

# Computer

07.20

## BLOCKCHAIN: FROM TECHNOLOGY TO MARKETPLACES

IEEE

IEEE  
COMPUTER  
SOCIETY

vol. 53 no. 7

[www.computer.org/computer](http://www.computer.org/computer)



# IEEE Computer Society Has You Covered!

**WORLD-CLASS CONFERENCES** — Stay ahead of the curve by attending one of our 200+ globally recognized conferences.

**DIGITAL LIBRARY** — Easily access over 780k articles covering world-class peer-reviewed content in the IEEE Computer Society Digital Library.

**CALLS FOR PAPERS** — Discover opportunities to write and present your ground-breaking accomplishments.

**EDUCATION** — Strengthen your resume with the IEEE Computer Society Course Catalog and its range of offerings.

**ADVANCE YOUR CAREER** — Search the new positions posted in the IEEE Computer Society Jobs Board.

**NETWORK** — Make connections that count by participating in local Region, Section, and Chapter activities.

**Explore all of the member benefits at [www.computer.org](http://www.computer.org) today!**



# Computer

## GUEST EDITORS' INTRODUCTION

### Blockchain: From Technology to Marketplaces

KARTHIK NANDAKUMAR, NALINI RATHA, SHARATH PANKANTI,  
ALEX PENTLAND, AND MAURICE HERLIHY

JULY 2020

FEATURES

20

Blockchain Architecture  
for Auditing Automation  
and Trust Building in  
Public Markets

SEAN CAO, LIN WILLIAM CONG, MENG HAN,  
QIXUAN HOU, AND BAOZHONG YANG

29

PharmaCrypt:  
Blockchain for Critical  
Pharmaceutical Industry  
to Counterfeit Drugs

NEETESH SAXENA, IEUAN THOMAS,  
PROSANTA GOPE, PETE BURNAP,  
AND NEERAJ KUMAR

45

Blockchain for Video  
Streaming:  
Opportunities,  
Challenges, and  
Open Issues

NABAJEET BARMAN, DEEPAK G. C.,  
AND MARIA G. MARTINI



## ABOUT THIS ISSUE BLOCKCHAIN: FROM TECHNOLOGY TO MARKETPLACES

*Research challenges  
in adopting blockchain  
technology into mainstream  
applications.*

## FEATURES CONTINUED

### 57 **Blockchain for E-Health-Care Systems: Easier Said Than Done**

SUJIT BISWAS, KASHIF SHARIF, FAN LI,  
AND SARAJU P. MOHANTY

## Departments

### 4 **Elsewhere in the CS**

## Membership News

### 100 **CS Connection**

### 102 **IEEE Computer Society Information**

## COLUMNS

### 7 **SPOTLIGHT ON TRANSACTIONS**

Is Artificial Intelligence  
Able to Help With Pain  
Assessment?

ELISABETH ANDRÉ

### 9 **50 & 25 YEARS AGO**

ERICH NEUHOLD

### 11 **COMPUTING THROUGH TIME**

ERGUN AKLEMAN

### 12 **EIC'S MESSAGE**

The "Patching"  
Mentality

JEFFREY VOAS

### 68 **BODY OF KNOWLEDGE**

A Prize, a Prediction,  
and a Drama

DAVID ALAN GRIER

### 71 **CYBER-PHYSICAL SYSTEMS**

Research Challenges  
for Heterogeneous  
Cyberphysical System Design

SHUVRA S. BHATTACHARYYA  
AND MARILYN C. WOLF

### 76 **EDUCATION**

Contactless U: Higher  
Education in the  
Postcoronavirus World

PHIL LAPLANTE

### 80 **AFTERSHOCK**

It's On: COVID-19,  
Risk Ecology, and  
Preparedness Tips

HAL BERGHEL,  
ROBERT N. CHARETTE,  
EDWARD G. HAPP, AND  
JOHN LESLIE KING

### 88 **CYBERTRUST**

21 Years of Distributed  
Denial-of-Service:  
Current State of Affairs

ERIC OSTERWEIL, ANGELOS  
STAVROU, AND LIXIA ZHANG

### 93 **OUT OF BAND**

A Collapsing  
Academy, Part 1

HAL BERGHEL

**Circulation:** *Computer* (ISSN 0018-9162) 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. IEEE Computer Society membership includes a subscription to *Computer* magazine.

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



## EDITOR IN CHIEF

**Jeffrey Voas**  
IEEE Fellow  
[j.voas@ieee.org](mailto:j.voas@ieee.org)

## ASSOCIATE EDITOR IN CHIEF

**Elisa Bertino**  
Purdue University  
[bertino@cs.purdue.edu](mailto:bertino@cs.purdue.edu)

## ASSOCIATE EDITOR IN CHIEF, COMPUTING PRACTICES

**Rohit Kapur**  
Synopsis  
[kapurfamily04@gmail.com](mailto:kapurfamily04@gmail.com)

## ASSOCIATE EDITOR IN CHIEF, PERSPECTIVES

**Jean-Marc Jézéquel**  
University of Rennes  
[jean-marc.jezequel@irisa.fr](mailto:jean-marc.jezequel@irisa.fr)  
**George K. Thiruvathukal**  
Loyola University Chicago  
[gkt@cs.luc.edu](mailto:gkt@cs.luc.edu)

## 2020 IEEE COMPUTER SOCIETY PRESIDENT

**Leila De Floriani**  
University of Maryland,  
College Park  
[dflo@umiacs.umd.edu](mailto:dflo@umiacs.umd.edu)

## AREA EDITORS

### CLOUD COMPUTING

**Schahram Dustdar**  
TU Wien

### COMPUTER ARCHITECTURES

**David H. Albonesi**  
Cornell University  
**Erik DeBenedictis**  
Zettaflops, LLC

### CYBER-PHYSICAL SYSTEMS

**Oleg Sokolsky**  
University of Pennsylvania

### CYBERSECURITY

**Rick Kuhn**  
NIST

### DIGITAL HEALTH

**Christopher Nugent**  
Ulster University

### EMBEDDED COMPUTING

**Marilyn Wolf**  
University of Nebraska

### HIGH-PERFORMANCE COMPUTING

**Vladimir Getov**  
University of Westminster

### INTERNET OF THINGS

**Michael Beigl**  
Karlsruhe Institute of Technology

### SECURITY AND PRIVACY

**John Viega**  
Capsule8

### SOCIAL-PHYSICAL-CYBER SYSTEMS

**Mike Hinchey**  
University of Limerick

### SOFTWARE ENGINEERING

**Phil Laplante**  
Pennsylvania State University

### VISION, VISUALIZATION, AND AUGMENTATION

**Mike J. Daily**  
HRL Laboratories

## COLUMN AND DEPARTMENT EDITORS

### AFTERSHOCK

**Hal Berghel**  
University of Nevada, Las Vegas

### Robert N. Charette

ITABHI Corporation  
**John L. King**  
University of Michigan

### BODY OF KNOWLEDGE

**David Alan Grier**  
Djaghe, LLC

### COMPUTING'S ECONOMICS

**Nir Kshetri**  
University of North Carolina-  
Greensboro

### COMPUTING THROUGH TIME

**Ergun Akleman**  
Texas A&M

### CYBER-PHYSICAL SYSTEMS

**Dimitrios Serpanos**  
University of Patras

### CYBERTRUST

**James Bret Michael**  
Naval Postgraduate School

### DATA

**Norita Ahmad**  
American University of Sharjah

### Preeti Chauhan

Google

### EDUCATION

**Irena Bojanova**  
NIST

### THE IOT CONNECTION

**Trevor Pering**  
Google

### IT INNOVATION

**Mark Campbell**  
Trace3

### OPEN SOURCE EXPANDED

**Dirk Riehle**  
University of Erlangen-Nuremberg

### OUT OF BAND

**Hal Berghel**  
University of Nevada, Las Vegas

### REBOOTING COMPUTING

**Erik DeBenedictis**  
Zettaflops, LLC

### SOFTWARE ENGINEERING

**Phil Laplante**  
Pennsylvania State University

### SPOTLIGHT ON TRANSACTIONS

**Ron Vetter**  
University of North Carolina  
Wilmington

### STANDARDS

**Forrest "Don" Wright**  
Standards Strategies, LLC

### WEB EDITOR

**Zeljko Obrenovic**  
Incision

### 50 & 25 YEARS AGO

**Erich Neuhold**  
University of Vienna

## ADVISORY PANEL

Doris L. Carver, Louisiana State University (EIC Emeritus)  
Carl K. Chang, Iowa State University (EIC Emeritus)  
Bob Colwell, Consultant  
Sumi Helal, University of Florida (EIC Emeritus)  
Bill Schilit, Google  
Ron Vetter, University of North Carolina Wilmington (EIC Emeritus)  
Alf Weaver, University of Virginia



## CS PUBLICATIONS BOARD

Fabrizio Lombardi (VP of Publications), Cristiana Bolchini, Javier Bruguera, Carl K. Chang, Fred Douglass, Charles Hansen, Shi-Min Hu, Antonio Rubio, Diomidis Spinellis, Stefano Zanero, Daniel Zeng

## MAGAZINE OPERATIONS COMMITTEE

Diomidis Spinellis (Chair), Lorena Barba, Irena Bojanova, Shu-Ching Chen, Gerardo Con Diaz, Lizy K. John, Marc Langheinrich, Torsten Moller, David Nicol, Ipek Ozkaya, George Pallis, VS Subrahmanian, Jeffrey Voas

## COMPUTER STAFF

### Senior Managing Editor

Geraldine Krohn-Taylor  
[g.krohn-taylor@ieee.org](mailto:g.krohn-taylor@ieee.org)

### Cover Design

Nanette Hoogslag

### Peer Review Administrator

[computer-ma@computer.org](mailto:computer-ma@computer.org)

### Publications Portfolio Manager

Carrie Clark

### Senior Advertising Coordinator

Debbie Sims

### Publisher

Robin Baldwin

### IEEE Computer Society

### Membership Director

Erik Berkowitz

### IEEE Computer Society Executive

Director  
Melissa Russell

## IEEE PUBLISHING OPERATIONS

### Senior Director, Publishing

Operations  
Dawn M. Melley

### Director, Editorial Services

Kevin Lisankie

### Director, Production Services

Peter M. Tuohy

### Associate Director,

Information Conversion  
and Editorial Support

Neelam Khinvasara

### Senior Art Director

Janet Dudar

Digital Object Identifier 10.1109/MC.2020.2990594

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 the IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). Copyright © 2020 IEEE. All rights reserved. IEEE prohibits discrimination, harassment, and bullying. For more information, visit [www.ieee.org/web/aboutus/whatis/policies/p9-26.html](http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html).





# ELSEWHERE IN THE CS

## Computer Highlights Society Magazines

The 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.

---

### **computing** in SCIENCE & ENGINEERING

#### **Exploratory Metamorphic Testing for Scientific Software**

Scientific model developers are able to verify and validate their software via metamorphic testing (MT), even when the expected output of a given test case is not readily available. The tenet is to check whether certain relations hold among the expected outputs of multiple related inputs. Contemporary approaches require that the relations be defined before tests. In this article from the March/April 2020 issue of *Computing in Science & Engineering*, the authors' experience shows that it is often straightforward to first define the multiple iterations of tests for performing continuous simulations and then keep multiple and even competing metamorphic relations open for investigating the testing-result patterns. The authors call this new approach exploratory MT, and they report their experience of applying it to detect bugs, mismatches, and constraints in automatically calibrating parameters for the U.S. Environmental Protection Agency's Storm Water Management Model.

---

### **IEEE** **Annals** of the History of Computing

#### **The Font Wars, Part 1**

The Font Wars were a decades-long competition in the computer industry for dominance in font technology, viewed

as a key success factor for personal computing platforms. The Font Wars spurred innovative scientific research into the small, nearly subliminal forms of the printed letters on which modern civilization was based, yet which had received little scrutiny outside the printing trades. More than a business episode, the Font Wars were above all a manifestation and translation of ideas—some modern, some ancient, some theoretical, and some practical—into computer software and hardware. At the heart of the Font Wars was a fundamental question: What is the best way to turn traditional printed letter forms into digital fonts for computer screens and printers? Answers to this question were researched, implemented, and launched into the marketplace, where their intense competition transformed the 500-year tradition of printing and publishing—placing the electronic literacy on the screens of billions of digital displays, computers, tablets, and smartphones around the world. Read more in this article from the January–March 2020 issue of *IEEE Annals of the History of Computing*.

---

### **IEEE** **Computer Graphics** and Applications

#### **Illustrating Changes in Time-Series Data With Data Video**

Understanding the changes of time series is a common task in many application domains. Converting time-series data into videos helps an audience with little or no background knowledge gain insights and deep impressions. It essentially integrates data visualizations and animations to present the evolution of data expressively. However, it remains challenging to create this kind of data video. First, it is difficult to efficiently detect important changes and include them in the video sequence. Existing methods require much manual effort to explore the data and find changes. Second, how these changes are emphasized in the videos is also worth studying. A video without emphasis will hinder an audience from noticing those important changes. This



article from the March/April 2020 issue of *IEEE Computer Graphics and Applications* presents an approach that extracts and visualizes important changes of a time series. Users can explore and modify these changes and apply visual effects on them. Case studies and user feedback demonstrate the effectiveness and usability of the approach.

---

## IEEE Intelligent Systems

### Research on Road Traffic Situation Awareness System Based on Image Big Data

Road traffic is an important component of the national economy and social life. Promoting intelligent and Informationization construction in the field of road traffic is conducive to the construction of smart cities and the formulation of macrostrategies and construction plans for urban traffic development. Aiming at the shortcomings of the current road traffic system, this article from the January/February 2020 issue of *IEEE Intelligent Systems*—on the basis of combining convolutional neural networks (CNNs), situational awareness, databases, and other technologies—takes the road traffic situational awareness system as its research object and analyzes the information collection, processing, and analysis process. CNNs, region-CNN (R-CNN), fast R-CNN, and faster R-CNN are used for vehicle class classification and location identification in road image big data. The deep CNN model based on road traffic image big data was further established, and the system requirements analysis and system framework design and implementation were carried out. Through the analysis and trial of actual cases, the results show the application effect of the realized road traffic situational awareness system, which provides a scientific reference and basis for the establishment of modern intelligent transportation system.

---

## IEEE Internet Computing

### Container NATs and Session-Oriented Standards: Friends or Foe?

This article from the November/December 2019 issue of *IEEE Internet Computing* highlights issues that arise when deploying network address translation middle-boxes through containers. The authors focus on Docker as the container technology of choice and present a thorough analysis of its networking model with special attention to the default bridge network driver that is used to implement network address translation functionality. They discuss some unexpected shortcomings and elaborate on

the suitability of containers for deploying services based on the Interactive Connectivity Establishment standard protocol. To support their findings, they present experiments that they conducted in a real-world operational environment, namely a WebRTC service based on the Janus media server.

---

## IEEE micro

### Compute Solution for Tesla's Full Self-Driving Computer

Tesla's full self-driving (FSD) computer is the world's first purpose-built computer for the highly demanding workloads of autonomous driving. It is based on a new system-on-a-chip that integrates industry-standard components, such as CPUs, ISPs, and GPUs, with custom neural network accelerators. The FSD computer is capable of processing up to 2,300 frames per second, which is a 21× improvement over Tesla's previous hardware and at a lower cost. When fully utilized, it enables a new level of safety and autonomy on the road. Read more in this article from the March/April 2020 issue of *IEEE Micro*.

---

## IEEE MultiMedia

### Metric Learning-Based Multimodal Audio-Visual Emotion Recognition

People express their emotions through multiple channels, such as visual and audio ones. Consequently, automatic emotion recognition can be significantly benefited by multimodal learning. Even though each modality exhibits unique characteristics, multimodal learning takes advantage of the complementary information of diverse modalities when measuring the same instance, resulting in enhanced understanding of emotions. Yet, their dependencies and relations are not fully exploited in audio-video emotion recognition. Furthermore, learning an effective metric through multimodality is a crucial goal for many applications in machine learning. Therefore, in this article from the January–March 2020 issue of *IEEE MultiMedia*, the authors propose multimodal emotion recognition metric learning, learned jointly to obtain a discriminative score and a robust representation in a latent space for both modalities. The learned metric is efficiently used through the radial basis function-based support vector machine kernel. The evaluation of the framework shows a significant performance, improving the state-of-the-art results on the eNTERFACE and CREMA-D data sets.





### CLIMB: A Pervasive Gameful Platform Promoting Child Independent Mobility

Child independent mobility (CIM) refers to the freedom and capability of children to move about their local neighborhoods without constant direct adult supervision. Our CLIMB project combats an observed decline in CIM, offering a pervasive gameful platform for home-school mobility composed of three primary components: the first two using technology to support different levels of child independence and the third providing an element of continuous motivation for positive behavior change. This article from the January–March 2020 issue of *IEEE Pervasive Computing* describes these three novel technologies: PedibusSmart, SafePath, and KidsGoGreen. It reports on four years of success with more than 1,800 elementary-age children, their teachers, and their families. The authors further show how 1) disappearing, pervasive technology contributes to successful adoption; 2) properly balancing trust and tracking leads to useful, noninvasive technological support; and 3) in-classroom, gameful technology engages and motivates participation, with behavior changes persisting over time.



### The Need for New Antiphishing Measures Against Spear-Phishing Attacks

In this article from the March/April 2020 issue of *IEEE Security & Privacy*, the authors provide extensive analysis of the unique characteristics of phishing and spear-phishing

attacks, argue that spear-phishing attacks cannot be well captured by current countermeasures, identify ways forward, and analyze an advanced spear-phishing campaign targeting white-collar workers in 32 countries.



### Three Phases of Transforming a Project-Based IT Company Into a Lean and Design-Led Digital Service Provider

Digital transformation requires a continuous review of value creation, value capture, and resourcing. In this article from the March/April 2020 issue of *IEEE Software*, the authors define a systematical service design concept to enable all stakeholders to achieve better outcomes in cocreation activities.



### Detecting Online Content Deception

The surge of deceptive content (such as fake news) in the past few years has made content deception an important area of research. The authors of this article from the March/April 2020 issue of *IT Professional* identify two main types of content deception based on either fake content or misleading content. They present a classification of deception attacks along with delivery methods. They also discuss defense measures that can detect deception attacks. Finally, they highlight some outstanding challenges in the area of content deception. ■

**Editorial:** Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *Computer* does not necessarily constitute endorsement by the IEEE or the IEEE 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 copyediting, proofreading, and formatting added by IEEE. For more information, please go to: [http://www.ieee.org/publications\\_standards/publications/rights/paperversionpolicy.html](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](mailto:pubs-permissions@ieee.org). Copyright © 2020 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 per-copy fee is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

A large background image showing several thick stacks of papers or documents, some with colorful edges, arranged in a way that creates a sense of depth and volume.

# Is Artificial Intelligence Able to Help With Pain Assessment?

Elisabeth André, Augsburg University

*This installment of Computer's series highlighting the work published in IEEE Computer Society journals comes from IEEE Transactions on Affective Computing.*

**M**ost people are able to communicate in an expressive manner when they are in pain. But what about people who are not able to report their pain experience or whose expression of pain is hard to interpret? Examples include individuals suffering from dementia, patients developing delirium, and newborns. Techniques that provide a reliable assessment of pain experience are a prerequisite for effective pain therapy. Due to recent advances in the automated detection and analysis of behavioral cues, the question arises of whether these techniques may help assess pain-related states in a reliable manner.

Digital Object Identifier 10.1109/MC.2020.2984873  
Date of current version: 1 July 2020

Physical pain is closely related to emotional states that may modulate the experience of pain and vice versa. Furthermore, principles and techniques from affective computing provide a solid basis for the automated analysis of pain-related states. Thus, it comes as no surprise that research on pain has increasingly attracted the interest of the affective computing community.

This trend is also reflected by the increased number of submissions to *IEEE Transactions on Affective Computing* that focus on technologies to detect and monitor pain.

The article "Automatic Recognition Methods Supporting Pain Assessment: A Survey,"<sup>1</sup> by Werner et al., presents the state of the art in automated pain recognition, focusing on facial expressions, body postures and movements, paralinguistic and linguistic vocalizations, and physiological signals, alone and in combination (Figure 1).

People show a great deal of individuality in their expression of pain, and there is no clear mapping between behavioral cues and the intensity and quality of pain. This even goes for experimental settings in which pain is induced in healthy people under controlled laboratory





**FIGURE 1.** Examples of facial expressions associated with pain.<sup>1</sup>

conditions. Thus, considerable effort has to be spent to establish a “gold standard” against which to evaluate the performance of pain detection components. Various instruments have been

developed to assess the experience of people in pain. The article describes clinically used pain assessment tools, such as self-reports and observational scales. It points out that representative

data are required for developing and validating techniques for pain detection. To accelerate progress in pain research, a number of (to be announced) publically available databases, some of which (BioVid, SenseEmotion, and EmoPain) have been introduced in earlier issues of *IEEE Transactions on Affective Computing* or are in early access, are presented.

To provide the reader with a realistic sense of the potential, but also the limitations, of automated pain detection, this survey article reviews more than 100 papers on this topic, obtained by searching the Web of Science as well as the proceedings of major conferences and journals on biomedical informatics and artificial intelligence (including their reference lists). Particular challenges arise due to variations in behavioral expressions that are only indirectly related to pain. This article provides guidelines on paths to take to overcome existing challenges. Promising directions of research include approaches to incorporate knowledge of the context in which pain is observed and studies on the interaction of physical pain with other affective states. **C**

## REFERENCE

1. P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. Picard, “Automatic recognition methods supporting pain assessment: A survey,” *IEEE Trans. Affect. Comput.*, to be published. doi: 10.1109/TAFFC.2019.2946774.

**ELISABETH ANDRÉ** is a full professor of computer science at Augsburg University, Germany, where she is the chair of the Human-Centered Multimedia lab, and editor-in-chief of *IEEE Transactions on Affective Computing*. Contact her at [andre@informatik.uni-augsburg.de](mailto:andre@informatik.uni-augsburg.de).

# 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.

**IEEE Software**

Digital Object Identifier 10.1109/MC.2020.3000398

Author guidelines:  
[www.computer.org/software/author](http://www.computer.org/software/author)  
 Further details: [software@computer.org](mailto:software@computer.org)  
[www.computer.org/software](http://www.computer.org/software)

# 50 & 25 YEARS AGO



**EDITOR ERICH NEUHOLD**  
University of Vienna  
erich.neuhold@univie.ac.at



## JULY 1970

We will be skipping July 1970, and the next time content from the 1970s will appear is in the September issue.

## JULY 1995

[www.computer.org/csdl/mags/co/1995/07/index.html](http://www.computer.org/csdl/mags/co/1995/07/index.html)

**The End of Work as We Know It** (p. 10) “In the United States alone, corporations are eliminating more than 2 million jobs annually. Men and women everywhere are worried about their future. ... He goes on to explain how the computer revolution has reduced the US manufacturing workforce from 33% in 1950 to 17% today. Meanwhile, from 1979 to 1992, productivity increased by 35%. This is a result of computerization, not foreign competition. ... We have entered the info age, where virtual worlds replace factories and telepresence replaces commuting. ... The hierarchical, monolithic, assembly-line manufacturing plant is a relic of the industrial age. This dinosaur is sinking in a swamp while the fleet-of-foot SOHO (small office-home office) is taking over.” (p. 11) “It’s not as difficult to prepare for the info age as it might sound. ... Whatever your formal training, don’t forget to keep your info age skills sharpened. In the future, expect to see the end of work, over and over again.” [Editor’s note: Looking at the United States and other developed countries, the manufacturing loss did happen. However, the article totally missed the coming of two things: first, the tremendous rise in the global economy, with huge manufacturing organizations in countries like Korea, China, and Vietnam, and second, the growth in service industries of all kinds, such as banking, logistics, sales, and support.]

**Guest Editors’ Introduction: Virtual Reality: In the Mind of the Beholder** (p. 17) “The elements: Interaction is the process of inputting data to the system and receiving data from it. The 3D graphics, a form of computer output, let users ‘see’ the virtual environment. Immersion refers to the user’s feeling of ‘presence’ in the virtual world. An immersive application convinces

users that they are in a replicated environment.” (p. 18) “Our playbill features medical imaging, psychological treatment, simulations, visualization, and terrain database construction. In contrast, the media have glorified frivolous VR applications, touting VR video games, ignoring serious research, and making outlandish assertions about VR as if these advances were already accomplished.” [Editor’s note: This view, of course, misses the fact that so-called frivolous applications have been driving the field for many years. Cost reductions and sophisticated hard- and software were the result of those developments.]

**Two-Handed Spatial Interface Tools for Neurosurgical Planning** (p. 20) “Neurosurgery is inherently a three-dimensional activity. It deals with complex structures in the brain and spine that overlap and interact in complicated ways.” (p. 22) “Therefore, in Netra, users manipulate virtual objects seen on a standard workstation monitor by moving the props with their hands. Since many people associate the phrase ‘virtual reality interface’ with immersing head-mounted displays, we often characterize our system as a *spatial desktop interface*—spatial because it involves moving six-degree-of-freedom input sensors in free space, desktop because it uses a standard monitor on the user’s desk.” (p. 25) “Proceeding from a presurgical plan to the actual patient in the operating room requires transforming the coordinate system *I* of the volumetric image data to the coordinate system *L* of Leksell space. This is accomplished by imaging the patient on the morning of surgery with a Leksell frame that has been attached to the patient’s head and fitted with a special fiducial system.” [Editor’s note: This is a very interesting article concerning the interaction between virtual objects and the real world; many of the things discussed have now moved into mainstream medicine, especially in the area of microinvasive treatments.]

**Virtual Environments for Treating the Fear of Heights** (p. 27) “Acrophobia, a simple phobia, is characterized by marked anxiety upon exposure to heights, by avoidance of heights, and by interference in functioning as a result of this fear. Behavioral therapy of acrophobia has included



exposing the subject to anxiety-producing stimuli while allowing this anxiety to attenuate.” (p. 28) “Building environments for therapy, we designed a number of virtual height situations to correspond to the types used for in vivo stimuli. ... We created three virtual environments for use in the therapy sessions: an elevator, a series of balconies, and a series of bridges.” (p. 32) “In summary, our controlled study of applying virtual reality to exposure therapy of acrophobia has yielded remarkable results. ... Subjects experienced a range of physical anxiety symptoms consistent with the apparent threat they encountered. The degree of anxiety and habituation observed would not have occurred unless the subjects felt present in height situations.” (p. 33) “We have documented evidence for the experience of a sense of presence in an immersive virtual environment. We have also shown that a person’s perceptions of physical-world situations and behavior in the physical world can be modified by experiences in a virtual world.” [Editor’s note: Around the same time, other applications, for example, fear of flying, led to similar results. Of course, using virtual reality (VR) for learning and training purposes has become mainstream during the time since 1995.]

**The Iowa Driving Simulator: An Immersive Research Environment** (p. 35) “This simulator’s rich, fully interactive environment provides varied scenarios for meeting experimental needs—for example, engineering evaluation of automated highway systems. ... The IDS immersive virtual environment represents the driving experience with a maximum degree of fidelity and realism. To achieve this, it provides a full range of sensory cues—visual, motional, auditory, and haptic—to the driver of the simulated vehicle. The driver is placed in full control of this vehicle, which is represented by a detailed, physics-based mathematical model.” (p. 40) “With support from ARPA and assistance from the Army Combat Systems Test Activity, an IDS virtual proving-ground environment that closely duplicates two test courses at the Army’s Aberdeen Proving Ground (APG) has been developed. ... The data from these experiments is still being analyzed, but initial results indicate a high degree of correlation between on-course and simulator data for basic driver-performance measures: vehicle speed versus position on course, steering behavior, and pedal use.” [Editor’s note: Since 1995, VR applications for learning and training have developed rapidly and moved into the mainstream, but they have also been used in other areas, for example, in technical and performance simulations in all kinds of planning systems.]

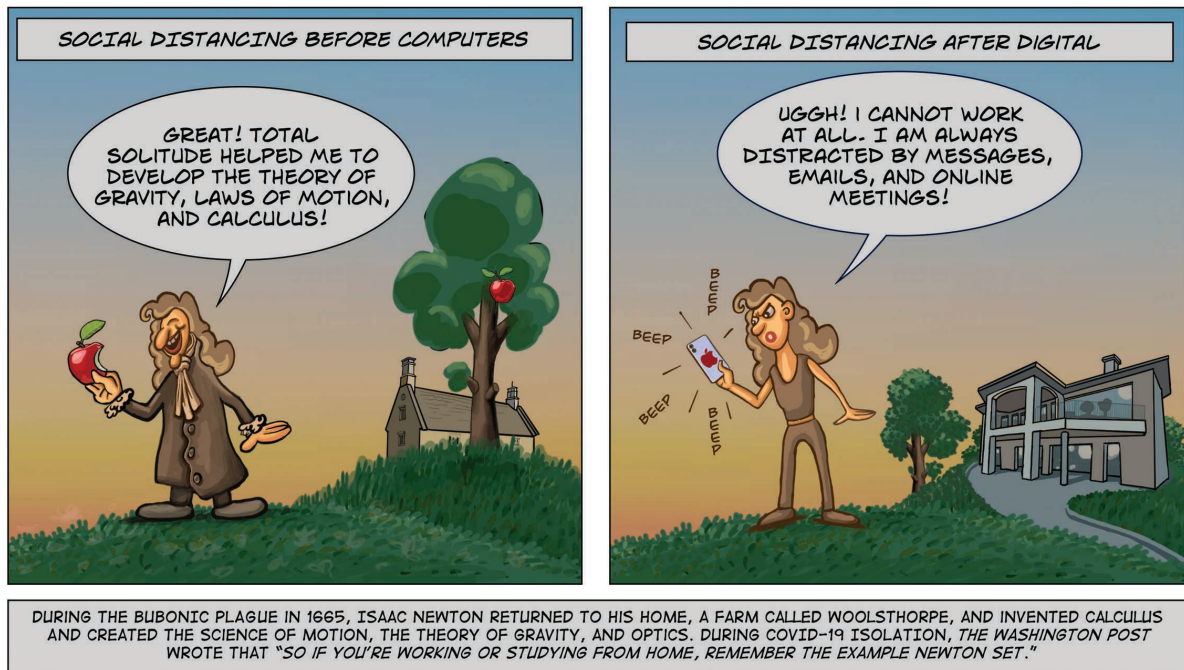
**The Responsive Workbench: A Virtual Work Environment** (p. 42) “Although as yet unrealized, the vision of an ultimate medium—using all of the human senses—is leading human-computer interface design toward virtual reality systems. ... These approaches aim for a universal interface (that is, intended for all users). ...” (p. 43) “Because we believe that responsive environments offer great potential for the human-computer interface, we developed the Responsive

Workbench as an alternative to other multimedia and virtual reality systems. ... The Responsive Workbench resulted from a joint effort of computer scientists, engineers, architects, and physicians to design a virtual environment. This article analyzes the working environments and behaviors of different users. ...” (p. 47) “Initial results showed that a virtual environment requires a high-resolution color display (at least 1280 × 1024 pixels) and real time rendering capability of complex objects with significant reflection and texture properties. The advantages offered engineers by the Responsive Workbench as a nonimmersive virtual environment compared with the BOOM system, for example, are the cooperative work setting and the incorporation of multisensory interaction models.” [Editor’s note: Nonimmersive as well as immersive VR systems have rapidly developed since this article was published. Again, like in other areas, the game industry has driven the development of many system components and led to competitive pricing structures.]

**A Large-Scale Complex Virtual Environment for Team Training** (p. 49) “Virtual environments that allow multiple participants to cooperatively interact present complex design problems. This Army program’s approach relies successfully on concurrent engineering, spiral development, and usability engineering. ... CCTT is a US Army program that when completed, will train ground combat tank and mechanized infantry forces on simulated equipment using a high-fidelity representation of actual terrain. Although CCTT is primarily a training system, its simulations will eventually be used for analytic studies, scientific experimentation, development activities, and engineering analyses. The CCTT system consists of networked simulators and workstations that collectively provide a virtual environment for training units to meet established Army standards.” (p. 55) “We are pioneering methods to achieve this, which may serve as prototypes, through spiral software development and periodic user evaluations.” [Editor’s note: This interesting article investigates, in detail, the pro and cons of systematic software development, in this case, a large virtual environment with 50 different user interfaces and half a million lines of code to be used for various U.S. Army applications.]

**Automating the Construction of Large-Scale Virtual Worlds** (p. 57) “Databases for large-scale virtual worlds have several critical applications. Automating their construction can improve fidelity and save considerable time. ... The focus of this article is the process by which the synthetic environment’s geographical component is constructed to model the real world with sufficient fidelity to support effective training and rehearsal. ... Many simulation projects require the data construction team to significantly augment standard Defense Mapping Agency (DMA) products to address critical issues of timeliness, local geographic intensification, and operational security.” (p. 63) “One key task in rapidly constructing virtual worlds is updating existing cartographic source material in a timely manner with new information extracted from imagery.” (p. 65) “This area

# COMPUTING THROUGH TIME



Digital Object Identifier 10.1109/MC.2020.2991275  
Date of current version: 1 July 2020

has much to gain from automated computer vision for mapping through stereo and monocular image analysis. At present, such methods still require manual correction. However, given appropriately high-resolution imagery, these methods offer a credible first pass at building detection and delineation. Continual improvements in such capabilities will enable population of virtual world databases quickly and cheaply.” [Editor’s note: The semiautomatic construction of virtual worlds as images of real ones plays an important role in road, building, and recreational-area construction. This article presents an analysis of the problems encountered. In part, they stem from the many inhomogeneous data collections that have to be integrated into one simulation basis.]

**National Productivity and Computers** (p. 66) “We are the first generation of Americans who think that our children will not live substantially better than we do. This article takes a hard look at US productivity, education, technology, and the prospects of improving national output. In the United States, national output has been growing very slowly over the past twenty years, and the average real wage has been stagnant. What little growth there has been is due to a growing work force and largely to the entry of women into the working world.” (p. 67) “The Rise of Science: The war’s end brought about another change that had both good and bad effects. The great and very visible achievements of scientists during the war—for example, the atomic bomb and radar—gave both politicians and the

public a feeling (and in my opinion a correct feeling) for the immense power that resides in scientific knowledge. And this thought—that science is power—led to a government emphasis on science and basic science support. ... As a consequence, manufacturing went its own way toward an eventual rude awakening. ... Education: Like science and advanced technology, education is something that the US turns to in moments of crisis. But it is often brought in as an explanation of more than it can explain. ... US schools below the college level are widely, and probably correctly, believed to have decayed. Their students certainly have test scores lagging behind those of many other advanced countries.” (p. 69) “Ways of Working: In the 1970s, I had the opportunity to visit Japan ... but much more important was the inherent excellence and rapidity of the Japanese development and manufacturing effort.” (p. 72) “And this is complicated by the possibility that the same skills you have may be available in a less developed country at a much lower price. And cheap sea transportation and cheapness of information transmission are rapidly making that competitive person into the person next door.” [Editor’s note: In this very interesting article, the author foresaw the transfer of manufacturing outside the United States but did not see that the service industry would rise to preplace, to a large degree, the jobs that got lost. However, it was correct that, at the same time, job security and increased earnings got lost. In that way, maybe, with the exception of the IT industry, we do not live better than our parents did.]



# The "Patching" Mentality

Jeffrey Voas, IEEE Fellow

*Software patches—we can't live without them, but it would be good if we could.*

**W**hen I was growing up, "patching" was something a person did to physical entities, such as clothing or pipes, whenever holes manifested. I also knew that cloth patches were used to sew quilts.

In the IT community, the term *software patch* has been employed to define software modifications (fixes) that mitigate code security vulnerabilities or other reliability/performance problems after software has been released. Generally speaking, patches are bandages applied after release and during maintenance, and they are often continuous over the lifetime of a product. A software patch should mitigate the situation in which there was an original weakness, or at some point, software behavioral decay occurred (possibly created by a modification of the operational environment).

Although this all sounds simple enough, the concept of software patching still remains somewhat mysterious to me. In fact, I'm not even sure if the term patch is used much anymore. After all, not all software vendors wish to acknowledge original weaknesses and defects. Instead, it seems that terms like *update* and *upgrade* are now the new,

preferred terms. So, have update and upgrade become replacement terms to move away from the uglier term of patch?

I've also always wondered about ideas, such as "building quality in" or "building security in." Is there a true business model for these concepts in a time-to-market world? After all, no one wants to be last to market. I believe that when it affects a brand, the answer is yes. But other than brand and stockholder protection, does it really matter? If the customer is willing to accept patch number 1, then patch number 2, and then patch number 3, and so on, why would an organization sacrifice time to market to offer excellence on day infinity (that is, the software product never gets released)? One might argue "yes" due to liability, but that argument is not strong. (And I don't even want to get into discussing patching patches.)

I ask these questions because I've been troubled for many years by an actual situation from 30 years ago in which a request for proposals was set up in such a way that the winning vendor was encouraged to bid on a job where the cost to perform the contract was greater than the revenue from the contract. So, who would bid on such a job? Well, as it turns out, many vendors because the profits were built into the maintenance phase to fix the problems built into the original deliverable. The lower the quality on the front end, the bigger the revenue/profits on the back end. That never sat well with me.



Patches are bandages applied after release and during maintenance, and they are often continuous over the lifetime of a product.

Automobile companies are an industry that understands the costs of recalls, not just in terms of fixing the vehicles that come back but in brand degradation. Does the software industry feel similar or even need to care?

To consider this question, just start from the simple fact that software is nonphysical. Physical is not easily malleable. Software is easily malleable. In

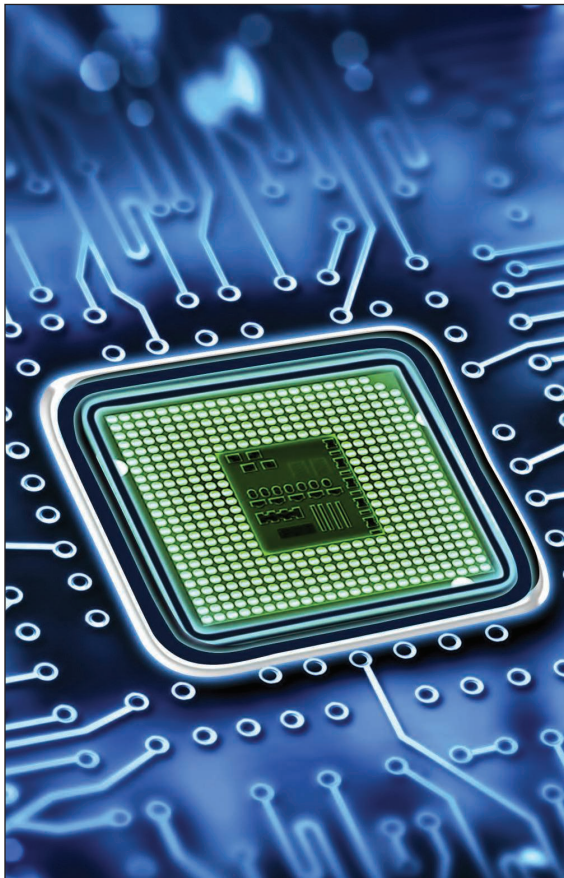
my opinion, software's malleability is the grand enabler of the "patching mentality." Whether good or bad, it is what it is, and it's here to stay.

In summary, installing patches, updates, upgrades, or whatever you prefer to call it is a normalized practice today. I believe we all agree that defect-free software is not an option. Zero trust, as an operational

philosophy, seems to be gaining momentum, but it has its own limits. In terms of patches to improve security, the release of new patches exacerbates successful attacks for those who do not install them immediately.

So, patches, we can't live without them, but it would be great if we could. It's a strange business model. Something to ponder. ■

**JEFFREY VOAS** is the editor in chief of *Computer*. He is a Fellow of the IEEE. Contact him at [j.voas@ieee.org](mailto:j.voas@ieee.org).



IEEE TRANSACTIONS ON

# COMPUTERS

## Call for Papers: *IEEE Transactions on Computers*

Publish your work in the IEEE Computer Society's flagship journal, *IEEE Transactions on Computers*. The journal seeks papers on everything from computer architecture and software systems to machine learning and quantum computing.

Learn about calls for papers  
and submission details at  
[www.computer.org/tc](http://www.computer.org/tc).





# Blockchain: From Technology to Marketplaces

**Karthik Nandakumar**, IBM Research

**Nalini Ratha and Sharath Pankanti**, IBM Thomas J. Watson Research Center

**Alex Pentland**, Massachusetts Institute of Technology

**Maurice Herlihy**, Brown University

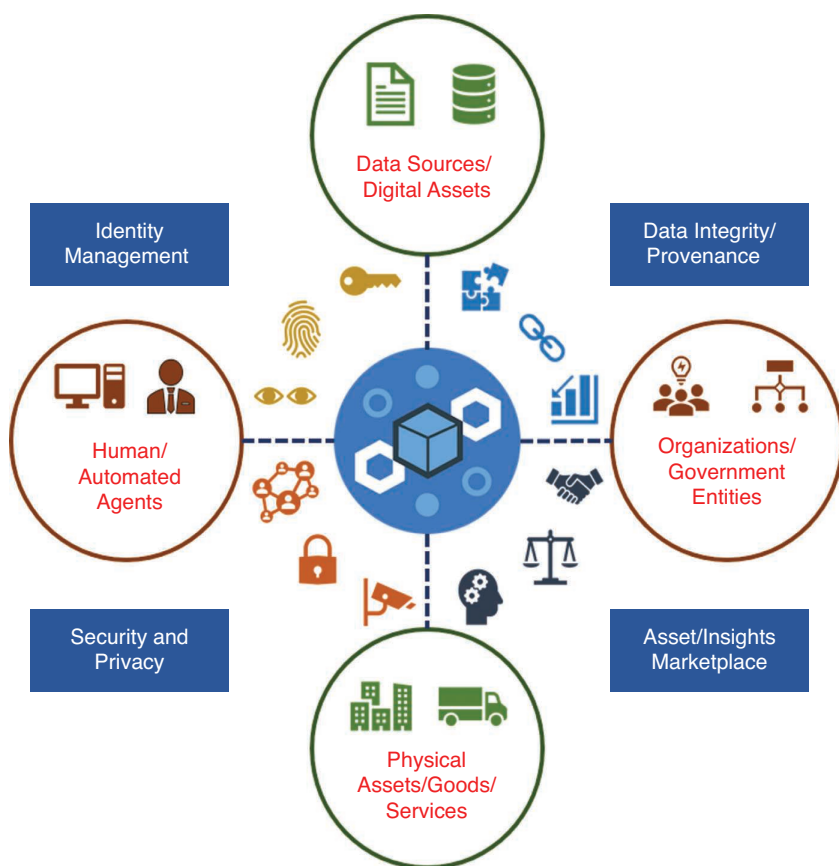


Digital Object Identifier 10.1109/MC.2020.2990776  
Date of current version: 1 July 2020

*This theme issue provides a glimpse of the diverse research challenges in adopting blockchain technology into mainstream applications. The four articles focus on the following core issues: scalability, transparency versus privacy, standardization, ecosystem, and integration.*

**T**rust and trust management lie at the heart of today's increasingly decentralized economy. In the past, trust has been enabled through a central authority. Blockchain, in all its variations, is emerging as a foundational technology that allows mutually untrusting parties to reach consensus on a shared digital history without a (central) trusted party. At the core of a blockchain application is a distributed immutable data store, and that is managed through smart contracts. Although blockchain is best known as the underlying core infrastructure of cryptocurrencies, it has many promising applications in other application domains, such as identity management, discovering critical obstacles in a complex supply chain, detecting money laundering and other financial crimes, identifying fake content, and better diagnoses of diseases (see Figure 1).

Several early adopters of blockchain are already reaping business benefits<sup>1</sup> by building solutions centered on trust, openness, and privacy. However, realizing the full potential of blockchain will require a significant level of fundamental advances in science and technology, major changes in business processes to create the right enabling environment, as well as innovative ideas to facilitate the integration of blockchain into real-world applications. This theme issue provides a glimpse of the diverse research challenges in adopting the technology into mainstream applications. Specifically, the four articles in this issue



**FIGURE 1.** Sample applications of blockchain. A wide variety of applications could significantly benefit from leveraging blockchain infrastructure,<sup>3</sup> which can address issues related to trust, governance, privacy, auditability, and provenance. Blockchain typically acts as a seamless distributed infrastructure that manages how (human or automated) agents as well as organizations interact with data sources/digital assets and physical assets under their control. Four broad classes of blockchain applications are illustrated. 1) Decentralized identity management could alleviate identity theft incidents<sup>4</sup> by prudently using various identifiers including biometrics. 2) Blockchain-based integrity and provenance can alleviate problems related to fake news<sup>5</sup> and other forms of disinformation. 3) Immutable audit trails made possible by blockchain could potentially enable a trusted marketplace for assets and insights.<sup>6</sup> 4) Increasingly adversarial cybersecurity incidents<sup>7</sup> could be addressed by auditable distributed ledgers.



focus on a mix of the following five core issues:

1. *Scalability*: One of the limitations of existing blockchain technologies is their inability to scale up to real-world, high-throughput applications without compromising on decentralization.
2. *Transparency versus privacy*: Another seemingly insurmountable issue is the tradeoff between transparency and privacy that is often encountered in blockchain applications.
3. *Standardization*: In a large family of applications, replacing a conventional centralized infrastructure with a blockchain infrastructure is not possible without standardizing the underlying data formats and interfaces.
4. *Ecosystem*: The efficacy of blockchain is maximized when the entire ecosystem of inter-related different applications could provide significantly better service by overhauling its consistent blockchain-compliant interfaces. Practitioners are recognizing that building an ecosystem is the most critical (and complex) effort for sustaining the benefits of blockchain infrastructure, and we increasingly are hearing the term minimal viable ecosystem in the context of blockchain infrastructures.<sup>2</sup>
5. *Integration*: Two distinct threads of evolution are emerging in the context of integrating blockchain into applications. The first approach can be referred to as transitioning,

which involves improving an application by gracefully transitioning from a conventional method with blockchain innovation so both can coexist in the transitional period. The second method is disruptive, which starts with recognizing a gap in a conventional application and bridging the limitation by replacing the conventional method with blockchain.

### IN THIS ISSUE

In “Blockchain Architecture for Auditing Automation and Trust Building in Public Markets,” Cao et al. show case how blockchain can enable the automated auditing of transactions as they occur by leveraging the strengths offered by the technology, such as immutability, load balancing, and differential access. This article also proposes a mechanism for privacy-preserving information exchange and highlights some of the scalability issues with existing blockchain protocols. Although their research exclusively focuses on auditing applications, the authors correctly note that similar advantages can be reaped by other applications by appropriately replacing their centralized infrastructures with blockchain.

“PharmaCrypt: Blockchain for Critical Pharmaceutical Industry to Counterfeit Drugs,” by Saxena et al., focuses on the problem of counterfeit drugs by first reviewing its widespread existence and the lack of solutions that can effectively ensure that both patients and dispensaries are made aware of the provenance of the drug. Their proposed solution describes a means of overhauling one of the most effective conventional centralized solutions based on

radio-frequency identification (RFID) technology and thus enabling a smooth transition to blockchain. They realistically conclude that their prototype is merely a new beginning toward an eventual solution, which requires building a viable ecosystem.

The article “Blockchain for Video Streaming: Opportunities, Challenges, and Open Issues,” by Barman et al., presents the role of blockchain technology in video streaming applications. The key message of this article is that a lack of standardization is critically debilitating the adoption of blockchain for media streaming. The authors propose a conceptual, unifying framework and interface for video streaming applications and observe the need for researchers to address several key technical challenges, such as scalability and privacy. They also stress the importance of appropriate business models to bring the technology to the marketplace successfully.

In the last feature article, “Blockchain for E-Health-Care Systems: Easier Said Than Done,” Biswas et al. give the readers a glimpse of the complexity of implementing a blockchain solution in the real world. The authors observe that health care is delivered through an ecosystem of closely connected networks of related interoperable services. Hence, the blockchain solutions implemented by individual health service providers must also be interoperable, which requires standards and new protocols for trade and consensus management. Another critical requirement in e-healthcare systems is the privacy of patient data. The authors conclude that to make a blockchain-based health-care solution a long-term success, it is critical to precisely capture and address its plethora of requirements.

## ABOUT THE AUTHORS

**KARTHIK NANDAKUMAR** is a research staff member at IBM Research, Singapore. His research interests include blockchain, computer vision, statistical pattern recognition, biometric authentication, image processing, and machine learning. Nandakumar received a Ph.D. in computer science from Michigan State University. He is a senior associate editor for *IEEE Transactions on Information Forensics and Security* and an associate editor for *Pattern Recognition*. He has received a number of awards including the Best Paper Award from *Pattern Recognition* (2005), Best Scientific Paper Award (Biometrics Track) at ICPR 2008, and the 2010 IEEE Signal Processing Society Young Author Best Paper Award. He is a Senior Member of the IEEE. Contact him at [nkarthik@sg.ibm.com](mailto:nkarthik@sg.ibm.com).

**NALINI RATHA** is a research staff member in artificial intelligence (AI) research at the IBM Thomas J. Watson Research Center. His areas of interest include computer vision, AI, biometrics and fairness, and trust in AI. Ratha received a Ph.D. in computer science from Michigan State University, East Lansing. He received the IEEE Biometrics Council Leadership Award in 2019. He has cochaired several workshops on the topics of blockchain, biometrics, fairness, and trust in AI and coedited several special issues of IEEE publications. He is a Fellow of the IEEE and International Association of Pattern Recognition and an Association for Computing Machinery distinguished scientist. Contact him at [ratha@us.ibm.com](mailto:ratha@us.ibm.com).

**SHARATH PANKANTI** is a research staff member in the Artificial Intelligence (AI) Department at the IBM Thomas J. Watson Research Center. His research interests focus on building scalable, fair, and trusted computer vision applications and their performance evaluation. Pankanti received a Ph.D. in computer science from Michigan State University. He has coedited theme issues of *Computer* on the topics of biometrics and cognitive computing, and he also cochaired workshops on the topics of blockchain, video summarization, fairness, and data efficiency. He is a Fellow of the IEEE, IAPR, OSA, and SPIE. Contact him at [sharat@us.ibm.com](mailto:sharat@us.ibm.com).

**ALEX PENTLAND** directs the Massachusetts Institute of Technology (MIT), Cambridge, Connection Science and helped create and direct the MIT Media Lab and the Media Lab Asia in India. He is one of the world's most cited computational scientists, serves on the board of the UN Foundation's Global Partnership for Sustainable Development Data, co-led the World Economic Forum discussion in Davos, Switzerland, that led to the European Union privacy regulation General Data Protection Regulation, and was central in forging the transparency and accountability mechanisms in the UN's Sustainable Development Goals. He is a member of the U.S. National Academies, and his most recent books are *Building the New Economy* (MIT Press), *Trusted Data* (MIT Press), and *Social Physics* (Penguin). He is a Member of the IEEE. Contact him at [sandy@media.mit.edu](mailto:sandy@media.mit.edu).

**MAURICE HERLIHY** is the An Wang Professor of Computer Science at Brown University. Herlihy received a Ph.D. in computer science from the Massachusetts Institute of Technology. He is the recipient of the 2003 Dijkstra Prize in Distributed Computing, the 2004 Gödel Prize in theoretical computer science, the 2008 ISCA Influential Paper Award, the 2012 Edsger W. Dijkstra Prize, and the 2013 Wallace McDowell Award. He received a 2012 Fulbright Distinguished Chair in the Natural Sciences and Engineering Lecturing Fellowship, and he is fellow of the Association for Computing Machinery, National Academy of Inventors, National Academy of Engineering, and National Academy of Arts and Sciences. Contact him at [mph@cs.brown.edu](mailto:mph@cs.brown.edu).




**T**hese four articles show that overcoming challenges such as scalability and privacy require core scientific and technological advancements in areas like distributed

### ACKNOWLEDGMENTS

Karthik Nandakumar, Nalini Ratha, and Sharath Pankanti would like to thank the members of the IBM Blockchain community for their support.

## A HERCULEAN EFFORT AT CONSENSUS BUILDING AMONG A DIVERSE SET OF STAKEHOLDERS IS REQUIRED.

computing and cryptography. For challenges such as standardization and ecosystem building, a herculean effort at consensus building among a diverse set of stakeholders is required. These advances will enable the transformation of blockchain from a niche technology for cryptocurrencies into a general-purpose technology capable of achieving unprecedented levels of transparency, accountability, and intelligence in the way we do business. We hope that this issue of *Computer* will serve as a valuable resource for the research community. Finally, we enjoyed guest editing this issue and would like to thank the reviewers for their time in shepherding these articles through the review process. Please feel free to contact us if you have any questions. 

### REFERENCES

1. "Forward together: Three ways blockchain explorers chart a new direction," IBM, Armonk, NY, May 2017. [Online]. Available: <https://www.ibm.com/services/insights/c-suite-study/blockchain>
2. E. Abebe et al., "Enabling enterprise blockchain interoperability with trusted data transfer (industry track)," in *Proc. 20th Int. Middleware Conf. Industrial Track*, Dec. 2019, pp. 29–35. doi: 10.1145/3366626.3368129.
3. D. D. F. Maesa and P. Mori, "Blockchain 3.0 applications survey," *J. Parallel Distrib. Comput.*, vol. 138, pp. 99–114, Apr. 2020. doi: 10.1016/j.jpdc.2019.12.019.
4. H. Weisbaum, "More than 1 million children were victims of ID theft last year," *NBC News*, June 21, 2018. [Online]. Available: <https://www.nbcnews.com/business/consumer/more-1-million-children-were-victims-id-theft-last-year-n885351>
5. F. Davey-Attlee and I. Soares, "Fake news," *CNN*. Accessed on: May 14, 2020. [Online]. Available: <https://money.cnn.com/interactive/media/the-macedonia-story/>
6. G. Singh and J. Levi, "MiPasa project and IBM Blockchain team on open data platform to support Covid-19 response," IBM, Armonk, NY, Mar. 27, 2020. [Online]. Available: <https://www.ibm.com/blogs/blockchain/2020/03/mipasa-project-and-ibm-blockchain-team-on-open-data-platform-to-support-covid-19-response/>
7. G. Belani, "5 cybersecurity threats to be aware of in 2020," *IEEE Computer Society*, Washington, D.C. Accessed on: May 14, 2020. [Online]. Available: <https://www.computer.org/publications/tech-news/trends/5-cybersecurity-threats-to-be-aware-of-in-2020>



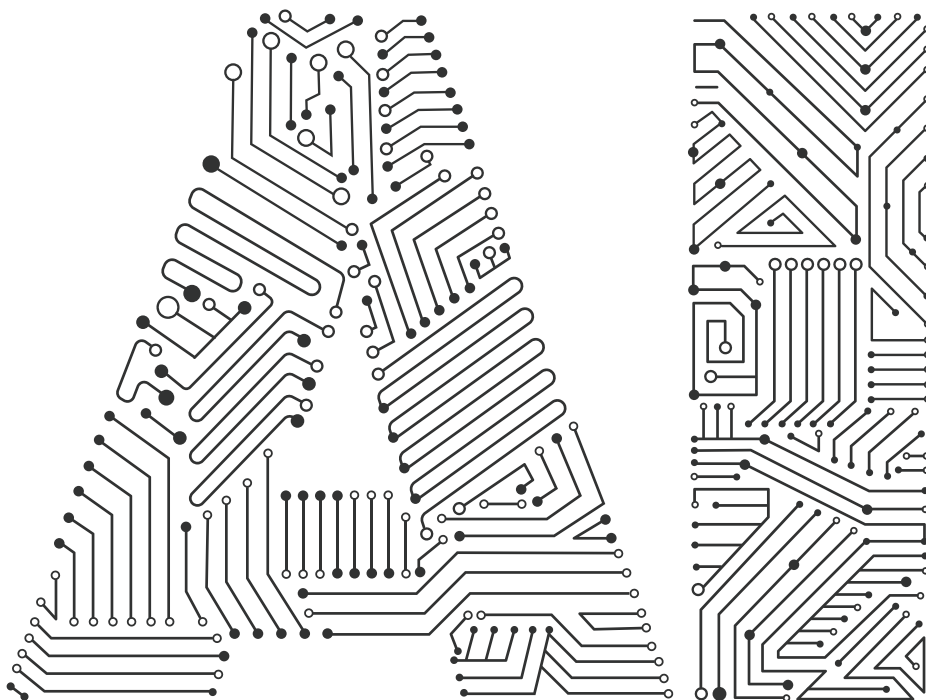
IEEE COMPUTER SOCIETY

**DIGITAL LIBRARY**

Access all your IEEE Computer Society subscriptions at

**computer.org**

**/mysubscriptions**



# AI'S 10 TO WATCH

## 2020 *IEEE Intelligent Systems* Magazine's "AI's 10 to Watch" CALL FOR NOMINATIONS

*IEEE Intelligent Systems* solicits nominations for its "Artificial Intelligence's 10 to Watch" in 2020. Those selected will appear in a special section of the magazine.

The 2020 nominees will be those who received their PhDs in the year 2014 or later and have already made significant contributions to one or more areas of AI.

A selection committee, chaired by a member of the *IEEE Intelligent Systems* Editorial Board, consists of prominent researchers from different subfields of AI and different geographic regions of the world.

Nominations from all areas AI are welcome.

For a list of past winners, and examples of the up-and-coming researchers we're looking for, please see 2018 winners here: [bit.ly/2TvG22y](https://bit.ly/2TvG22y)

### Nominations should include the following:

1. Nominee's CV and publication list
2. Two reference letters
3. Statement summarizing the candidate's achievements in AI (200 words maximum)

Nomination deadline:  
**10 July 2020**

Send nominations to  
[ieee@ten-to-watch-in-ai.com](mailto:ieee@ten-to-watch-in-ai.com)

Digital Object Identifier 10.1109/MC.2020.3000400



# Blockchain Architecture for Auditing Automation and Trust Building in Public Markets

**Sean Cao**, Georgia State University

**Lin William Cong**, Cornell University

**Meng Han**, Kennesaw State University

**Qixuan Hou**, Georgia Institute of Technology

**Baozhong Yang**, Georgia State University

*Business transactions by public firms must be reported, verified, and audited periodically, which is labor intensive and time consuming. To streamline this procedure, we have designed Future Auditing Blockchain to automate the reporting and auditing process, allowing auditors to focus on discretionary accounts to better detect and prevent fraud.*

**F**inancial auditing is a systematic and independent process of examining an organization's financial data, including books, accounts, statutory records, documents, and vouchers, to

determine if they are accurate and compliant with laws and regulations. Verification of counterparty transactions is an essential part of auditing. Public firms tend to be large, with a total global market capitalization of US\$68.7 trillion. Auditing firms handle large quantities of mechanical transaction verification and have limited resources for more sophisticated tasks that require

Digital Object Identifier 10.1109/MC.2020.2989789  
Date of current version: 1 July 2020



discretionary judgment and expertise. Due to the high cost of verification, auditors usually randomize audit samples. Consequently, traditional auditing is necessarily partial with a considerable potential for misreporting, which erodes investors' trust in public markets.<sup>1</sup>

Moreover, auditing firms may possess mutually useful information yet prefer to work independently because 1) clients are reluctant to authorize the sharing of data, which makes it illegal for third parties to do so, especially after regulations such as the General Data Protection Regulation; and 2) traditional infrastructure does not have a mechanism to share data in a cost-efficient way. In practice, when verifying transactions, auditors contact the transaction counterparties either manually or through a third party, which may not always be reliable.<sup>1</sup> Collaboration between auditing firms is challenging, primarily due to the lack of a system that is not only secure from hacking but also scalable and efficient in handling a large user base and multitudinous transactions.

As a potential solution to these problems, in this article we present Future Auditing Blockchain (FutureAB), a blockchain-based platform for collaborative auditing with advanced privacy protections. Thanks to the decentralized nature of blockchain, FutureAB can automate transaction auditing between firms without the need for a trusted third party. To ensure the privacy of proprietary data, we have adapted the Pedersen commitment to produce a modified data exchange scheme for detailed transfer of information along with the transactions. FutureAB also employs a smart wallet system and smart contracts to further

improve efficiency. We strengthened the protocol with ledgers to keep track of records with immutability and ensure informational security. Information stored on various ledgers makes it simple and easy to detect manipulation attempts. Finally, we implemented our FutureAB system on Ethereum to evaluate its performance. We found that FutureAB is scalable and efficient, with an encryption speed of 0.012 s/transaction and verification at 0.001 s/transaction.

Our system answers the auditing industry's call for blockchain-based innovation. Although all of the Big Four auditing firms are aware of the importance of blockchain and devoting vast resources to its development by establishing research labs or providing blockchain services,<sup>2,3</sup> it is still unclear exactly how this emerging technology will affect the auditing industry and indeed the auditors themselves. While accounting firms' recent efforts center on building in-house blockchain capabilities and services,<sup>2,4</sup> our work demonstrates the possibility of connecting isolated auditing processes while preserving data privacy with blockchain technology.

## RELATED WORK

### Collaborative and Continuous Auditing

In auditing, collaboration is often identified as a way to reduce costs and improve efficiency. There are several existing applications designed for collaborative auditing. Wu et al.<sup>5</sup> have proposed an agent-based architecture to increase the frequency of periodic audits. This scheme emphasizes efficiency in continuous auditing but does not address privacy concerns. Sachar et al.<sup>6</sup> have presented

a framework based on the concept of an audit warehouse that enables central, tool-supported auditing of cross-enterprise business processes. Chen et al.<sup>7</sup> have developed a collaborative continuous-auditing model relying on XML and Web Service technologies under service-oriented architecture environments. A complex protection profile is required to ensure data security in these two frameworks. Wang et al.<sup>8</sup> have proposed a secure cloud storage system to support secure public auditing and introduce a third party to check the integrity of data. However, the integrity and reliability of the third party are not guaranteed. Our blockchain architecture circumvents the aforementioned issues by implementing a decentralized verification mechanism.

### Smart Contract

Nick Szabo has proposed smart contracts, computer protocols that can automatically execute the terms of a contract, facilitate and verify the performance of a contract, interact with other contracts, make decisions, store data, and send data to others.<sup>13</sup> Many smart-contract platforms are now emerging, including Ethereum, Hyperledger, and Corda. The FutureAB platform takes full advantage of smart contracts to further minimize human error and improve efficiency in auditing.

### Commitment Schemes

The Pedersen commitment is a commitment scheme based on cryptographic hash functions.<sup>9</sup> A commitment scheme allows the sender to commit to a choice while hiding its selections from other receivers. Commitment schemes are widely used in

blockchain applications to preserve privacy. There are several examples of this tactic in recent literature. For example, Knirsch et al.<sup>10</sup> have proposed using commitment schemes in electronic vehicle charging; Zhang et al.<sup>11</sup> have proposed BCPay, a blockchain-based fair-payment framework for outsourcing services in cloud computing; and Xu et al.<sup>12</sup> have discussed the potential of commitment schemes for enabling sharing economies.

FutureAB leverages the Pedersen commitment with specific adjustments to guarantee information security. Specifically, to protect the participants' data privacy, our proposed application ensures the suppression of auxiliary information not directly related to transactions. For example, if company A in Atlanta shipped 1,000 units to company B in Phoenix, which arrived on 1 January 2020, the transaction record would contain affiliated information that might be potentially useful to competitors, such as date and type of product. Our goal is to provide not only a platform for the auditing process but also a mechanism that prevents the disclosure of unnecessary information, which can provide an incentive for intercompany auditor collaboration.

## SYSTEM DESIGN

### Pain Points of Current Auditing Processes

Traditionally, the auditing process of each company is independent. Several issues arise.

- › *High cost and low efficiency:* Auditors of one company must request transaction records from counterparties and manually verify the information, which is a labor-intensive process.
- › *Failure to fully utilize all information:* Reducing auditing sample size is a common way to reduce costs. However, Cao et al.<sup>1</sup> underscore that the sample size correlates with the quality of auditing; the failure of full information utilization therefore negatively affects the end result.
- › *Fraudulent reports:* Failure to use all information in auditing also creates a greater potential for fraud or misreporting. Companies may overstate earnings to boost their stock-market valuation.
- › *Privacy and access:* A platform for auditors to share transaction information might reduce the cost and improve the efficiency and quality

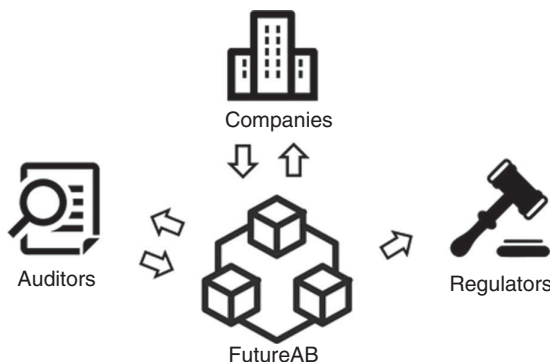
of the auditing process. However, companies are reluctant to reveal proprietary information to others, especially their competitors.

### Business Process Design

FutureAB addresses the aforementioned challenges. The platform focuses on auditing transaction-based accounts and, as shown in Figure 1, assists auditors in investigating mismatched transactions, companies being audited, and regulators who oversee these processes. The whole system is permission based, meaning that permission could be granted by the committee of the participants, such as the auditing association or the Company Public Accounting Oversight Board (PCAOB). The public key or address used within the blockchain would still be limited to the members of the system and would not be "public." All historical transactions are stored locally in the auditor members' proprietary databases.

Figure 2 shows the three major components of FutureAB.

- › *Private wallet:* Under the proposed architecture, each company possesses a private electronic wallet, called ABWallet, that stores its public addresses, private keys, and confidential information for encryption within the system. We next discuss how the wallet is generated, how addresses and keys are managed, and how the information is stored.
- › *Web service:* Auditors and regulators can access a public web-based application to perform tasks such as reviewing mismatched transactions. Smart contracts are deployed here to pair posted transactions for verification and then



**FIGURE 1.** Different roles in FutureAB.

write verified transactions onto the blockchain.

- ▶ **Blockchain:** Any key holder can use its private key to sign the verified transactions. The resulting signature is then recorded on the blockchain for peers to verify.

The business process of FutureAB is as follows.

1. **Initialize ABWallet for incoming companies:** A company can join the system by requesting access via the website. Once the access is granted, the incoming company can download ABWallet, which generates, stores, and manages public addresses, private keys, and commitment secrets for further activities on FutureAB. The organizations overseeing the auditing should be the ones who monitor the blockchain system and respond to the companies' access requests. As mentioned earlier, the PCAOB is a good candidate to maintain the auditing blockchain. An alternative would be an alliance of major auditing firms. The wallet initialization procedure is presented in Figure 3. When  $company_x$  joins the system, the company selects a set of other companies it often works with. In response, ABWallet generates distinct public addresses, private keys, and commitment secrets for the selected companies; sends public addresses and commitment secrets to the corresponding companies; and then requests addresses from the counterparty,  $company_y$ . FutureAB provides a company

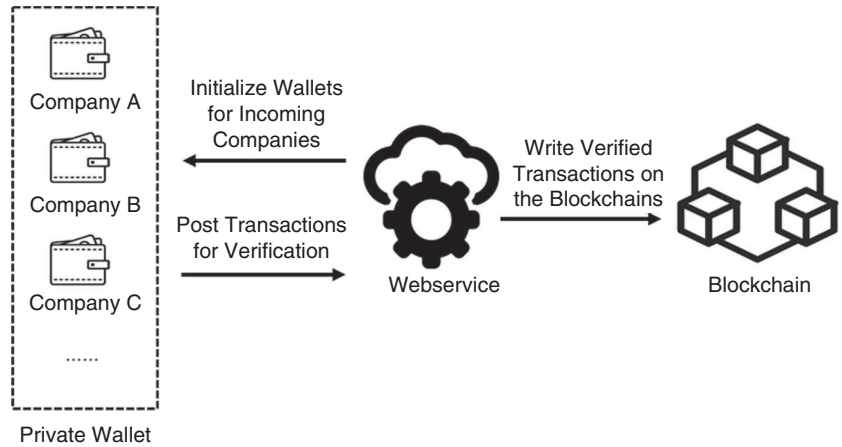


FIGURE 2. The business process design of FutureAB.

#### Wallet Initialization Procedure

```

all_companies  $\leftarrow$  all companies in the system
addressxy  $\leftarrow$  address of companyx for transactions with companyy
keyxy  $\leftarrow$  private key of companyx for transactions with companyy
secretxy  $\leftarrow$  secret to encrypt the amounts of transactions between
               companyx and companyy
G  $\leftarrow$  parameter of the elliptic curve used for ECDSA
Timestamp  $\leftarrow$  timestamp of storing the value set

Init_ABWallet (companyx):
  for (companyi in all_companies):
    secretxi  $\leftarrow$  random_generator ()
    keyxi  $\leftarrow$  random_generator ()
    addressxi  $\leftarrow$  keyxi * G
    send secretxi, addressxi, timestamp to companyi
    request addressix from companyi
    wait until (addressix is received):
      privately store the value set (timestamp,
                                   addressxi, addressix, keyxi, secretxi)

```

FIGURE 3. The wallet initialization procedure.

gallery, which sorts companies into different categories, such as a set of media companies, a set of healthcare companies, and a set of retail companies. The incoming company can pick several sets as its potential counterparties. This feature enables the wallet initialization

process to be more efficiently achieved in small batches. When posting transactions, ABWallet can automatically generate addresses and request counterparties' addresses if one of the counterparties does not yet exist in the wallet. Once the addresses are received, value



sets (timestamp,  $address_{xy}$ ,  $address_{yx}$ ,  $secret_{xy}$ ,  $key_{xy}$ ) are stored privately in the wallet of  $company_x$ . At the same time,  $company_y$  also privately stores the value sets (timestamp,  $address_{yx}$ ,  $address_{xy}$ ,  $secret_{xy}$ ,  $key_{yx}$ ). We use the Elliptic Curve Digital Signature Algorithm (ECDSA) as our signature scheme.

2. *Post transactions for verification:* ABWallet generates signatures and associated signed messages so that posting transactions on the web service will not compromise digital security. The post transaction procedure is presented in Figure 4. The messages are structured

as the sender address, receiver address, amount, date, and 0 or 1. 0 indicates the posted transaction is from the sender; otherwise, the last digit of the message is 1. Once the message is posted successfully, the status of the message is labeled as *pending*, which means that it is pending verification. Beyond basic information, companies are encouraged to post details of a transaction and hide the information from the public with a Pedersen commitment. If a discrepancy is spotted and auditors are notified, the auditors can request the commitments to be opened in order to review the details of the transactions. The auditors can

also contact the corresponding companies if more information is needed for the investigation.

3. *Verify posted transactions:* Both counterparties should post the transaction and encrypted messages with the same sender address, receiver address, and commitment secret. If there are multiple transactions between two companies within the same day (based on GMT), FutureAB will take the sum of these so that there is only one transaction between two companies per day. The web service consistently attempts to pair up two messages. A *pair* is defined as two messages with the same sender address, receiver address, and date. As described above, the last digit should be 0 or 1.

There are three possible states of one posted message, as shown in Figure 5.

- ▶ *Verified:* If the message is paired with another and two messages are identical except for their respective last digits, the transaction in the message is verified by both involved parties and can be written on the blockchain as a permanent record.
- ▶ *Risk:* A discrepancy is identified when a pair of messages contain different amounts. We label the pair as a *risk* to notify auditors to trigger an investigation. Being freed from mechanical transaction verification, auditors can focus on discretionary accounts where knowledge is indispensable.
- ▶ *Pending:* If only one involved party posts the transactions,

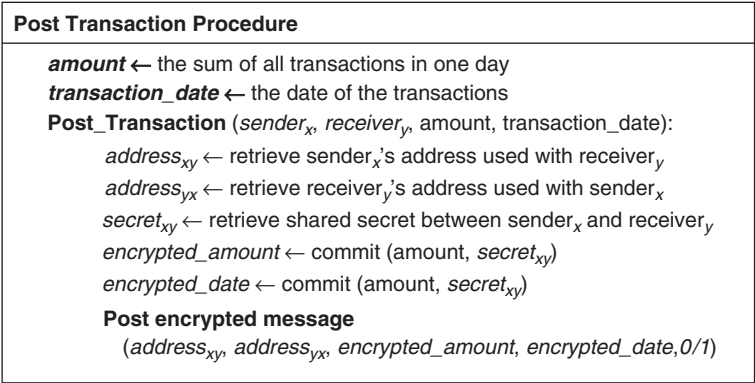


FIGURE 4. The post transaction procedure.

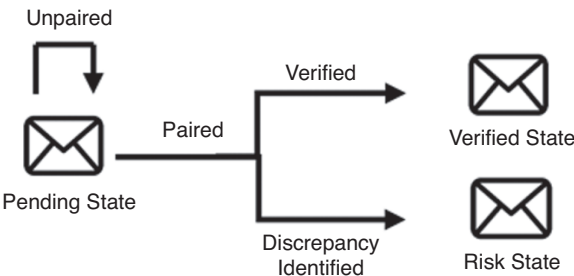


FIGURE 5. Three possible states of a posted message.

this asymmetry may produce messages that cannot be paired.

## TECHNICAL DETAILS

### ABWallet

We introduced ABWallet to allow each company to generate and manage value sets. ABWallet is also responsible for communicating the latest public addresses and commitment secrets between companies to ensure the synchronization of information. Whenever a company initiates the process of posting transactions, ABWallet retrieves the latest value sets and encrypts the transactions.

A member joins the system by downloading ABWallet and starting the wallet initialization procedure discussed in the previous section. Access to ABWallet should be kept private and secure. The only information flowing among different members' wallets should be public addresses and commitment secrets of counterparties, and the only information flowing from

a wallet to the web service should be signed messages.

For FutureAB, we propose using different addresses for transactions with different companies to preserve anonymity. We also recommend frequently generating new addresses to hide companies' identities in the posted transactions.

As shown in Figure 6, each wallet of the corresponding company is a row. For instance, company A has access to only the first row in the table. When company B generates a new value set for transactions with company A and updates the first cell in the second row, company A will be notified with new  $address_{BA}$  and  $secret_{BA}$ , and a new value set will be appended to the second cell in the first row. ABWallet helps companies generate value sets and maintain up-to-date information so that verification procedures can be executed quickly and correctly.

### Smart Contracts

Once smart contracts are compiled and migrated, the web service can implement them when certain conditions are

satisfied. In our setup, the smart contract is triggered when a new message is posted. It is executed to pair messages and then write verified messages on the blockchain. Compared to a traditional auditing system, smart contracts considerably reduce manual effort and costs in verifying transactions because they are code based and run live on the Internet at a low cost.

### Commitment Schemes

In FutureAB, only two involved parties share the secret to decrypt the message, meaning that the message is hidden from all other parties on the blockchain. In the meantime, both participating parties use the same secret to execute a transaction, which is then posted on the web service, indicating that both parties can no longer change what is committed. Note that FutureAB can accommodate transactions involving multiple parties.

This effort helps preserve the integrity of the content without disclosure; when inquiries for details are received,

	Company A	Company B	Company C
Company A		<div>timestamp_0 address_AB address_BA key_AB secret_AB</div> <div>timestamp_1 address_AB address_BA key_AB secret_AB</div>	<div>timestamp_0 address_AC address_CA key_AC secret_AC</div>
Company B	<div>timestamp_0 address_BA address_AB key_BA secret_AB</div> <div>timestamp_1 address_BA address_AB key_BA secret_AB</div>		<div>timestamp_0 ... timestamp_1 ...</div> <div>timestamp_3 ... timestamp_2 ...</div>
Company C	<div>timestamp_0 address_CA address_AC key_CA secret_AC</div>	<div>timestamp_0 ... timestamp_1 ...</div> <div>timestamp_3 ... timestamp_2 ...</div>	
Company X	...	...	...

FIGURE 6. The information stored in ABWallet.

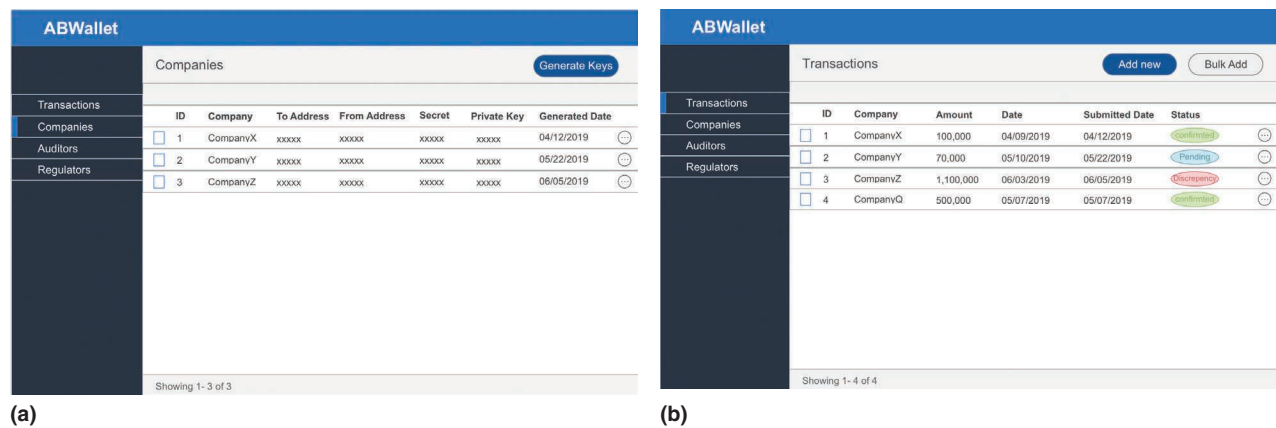


FIGURE 7. The user interface design of ABWallet.

the commitment can guarantee the trustworthiness of the committed content. We adapted the Pedersen commitment scheme in the design of FutureAB. The Pedersen commitments have the hiding property, which indicates that the commitment reveals nothing about the message. Additionally, Pedersen commitments are homomorphic, which facilitates the quick generation and verification of transactions on FutureAB by making it possible to combine commitments. If  $cm_1$  and  $cm_2$  are two commitments to values  $v_1$  and  $v_2$ , using commitment randomness  $r_1$  and  $r_2$ , then  $cm := cm_1 * cm_2$  is a commitment to  $v_1 + v_2$  using randomness  $r_1 + r_2$ . The commitment can preserve the security of certain information related to the transaction, such as transaction descriptions, quantities of products, and exchanged strategies, to the highest extent, which should encourage a larger number of participants to join and collaborate on the system.

IMPLEMENTATION DETAILS

Development Environment

The main components of FutureAB are ABWallet, the web application, and the

blockchain. The Ethereum blockchain is a public blockchain that allows users to perform Turing-complete calculations (smart contracts). The Ethereum protocol has an average block time of 15 s and charges small transaction fees for the processing of smart contracts. Ethereum satisfies all the current requirements of FutureAB. The confirmation time and other specific requirements of FutureAB should be the subject of future research.

The web application interacts with smart contracts on the Ethereum blockchain. The web app is written in JavaScript and HTML5, using the Truffle development framework. The framework enables JavaScript bindings for the smart contract and includes libraries such as web3.js that facilitate communication between the web app and the Ethereum client.

Rewards Program

FutureAB is designed to allow an optional rewards programs implemented to attract more companies and auditors. A rewards program can be introduced to motivate all parties in the system to actively and continuously post their transactions to

achieve collaborative and continuous auditing as an expected outcome of FutureAB. The rewards program can also motivate companies to mine and sign the verified transactions so that the records can be permanently written on the blockchain.

EVALUATION

To the best of our knowledge, FutureAB is one of the first platforms that can support collaborative and continuous auditing on the blockchain without compromising data privacy.

Once ABWallet is downloaded and initialized for  $company_x$ , users from  $company_x$  can review and manage partner companies' addresses, keys, and secrets via the list view shown in Figure 7. The context menu on each row allows users to generate new sets of values, to request new addresses from their counterparties, and to view the transaction history. Users can also view transaction details in the list view. The Bulk add button on the top right of the screen allows users to upload multiple transactions with an Excel sheet template. The status of each transaction is indicated in the status column.



We ran a simulation of wallet initialization and the transaction encryption on a 2.2-GHz Intel Core i7 machine with 16 GB of RAM and 1,600-MHz CPUs. This hardware is used only for simulation and performance evaluation. Since FutureAB is a distributed blockchain, the communication traffic and resource usage for each node are much lower than those on a single machine simulation once deployed. It takes 0.096 s on average to set up one value set for one counterparty. It takes about 16 min to set up the wallet for a new company when there are 10,000 selected counterparties. It takes 0.021 s on average to encrypt one transaction. Fewer than 7 min are needed to encrypt 20,000 transactions. As shown in Figure 8, the current system takes less than one minute to verify 10,000 transactions. As such, we believe that FutureAB can support real-time posting, that is, companies encrypting and posting transactions simultaneously with, or a short period of time after, the occurrence of the events.

In this article, we have presented a blockchain architecture to automate collaborative and continuous auditing and, in so doing, to build trust in public markets. Blockchain is one of the most influential emerging technologies in the past decade. The distributed nature of blockchains achieves peer-to-peer communication and allows auditing collaboration and financial reporting without relying on a trusted third party (decentralization). Blockchains naturally provide immutability, which guarantees that once an accounting activity is recorded, no one, including the owner of the business, can arbitrarily change the records (immutability). Moreover,

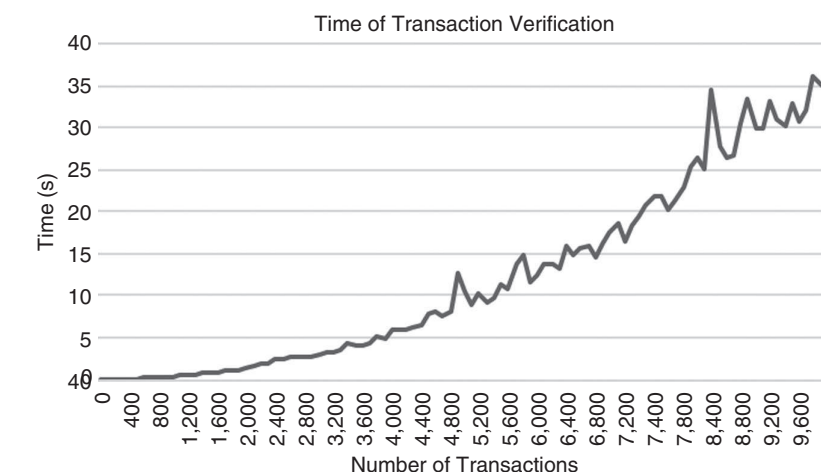



FIGURE 8. The time of transaction verification.

smart contracts use protocols and algorithms to digitally and automatically facilitate, verify, or perform a contract between two parties within a blockchain (automation). The use of encryption techniques protects proprietary information while ensuring certain messages can be recorded on a public blockchain without compromising privacy (encryption).

Although blockchains with encryption boast many exciting features, there are still several limitations. The transaction verification needs to pass a certain level of synchronization on the whole peer-to-peer system, which may result in a delay of seconds to minutes, not to mention the extra storage required. Fortunately, the rapid developments of hardware, consensus algorithms, and storage technology mitigate these concerns, especially for financial reporting and auditing processes that do not require a millisecond-level performance.

More generally, our proposed architecture provides an alternative way to achieve privacy-preserved information exchange. Besides auditing,

the system design could be applied to many other fields for information collaboration, including banking, insurance, and even health care. 

## ACKNOWLEDGMENTS

We would like to thank Jasmine Cheng and Fahad Saleh for insightful feedback and comments.

## REFERENCES

1. S. Cao, L. W. Cong, and B. Yang, "Financial reporting and blockchains: Audit pricing, misstatements, and regulation," *SSRN Electron. J.*, pp. 1–57, Sept. 25, 2018. doi: 10.2139/ssrn.3248002.
2. P. Bajpai, "Big 4 accounting firms are experimenting with blockchain and Bitcoin," *Nasdaq*, July 5, 2017. [Online]. Available: <https://www.nasdaq.com/articles/big-4-accounting-firms-are-experimenting-blockchain-and-bitcoin-2017-07-05>
3. A. Vetter, "Blockchain is already changing accounting," *Accounting Today*, May 7, 2018. [Online]. Available: <https://www.accountingtoday>

## ABOUT THE AUTHORS

**SEAN CAO** is an assistant professor at the School of Accountancy, Robinson School of Business, Georgia State University, Atlanta. His research interests include FinTech applications in capital markets. Cao received a Ph.D. in accountancy from the University of Illinois at Urbana-Champaign. Contact him at [scao@gsu.edu](mailto:scao@gsu.edu).

**LIN WILLIAM CONG** is the Rudd Family Professor of Management and associate professor of finance at Cornell University, Ithaca, New York, where he directs the FinTech Initiative. His research interests include economic data science, FinTech, and information economics. Cong received a Ph.D. in finance from Stanford University, California. Contact him at [will.cong@cornell.edu](mailto:will.cong@cornell.edu).

**MENG HAN** is an assistant professor in the College of Computing and Software Engineering at Kennesaw State University, Georgia. His research interests include data-driven intelligence, FinTech, and blockchain technologies. Han received a Ph.D. in computer science from Georgia State University, Atlanta. He is a member of the IEEE and IEEE Communications Society. Contact him at [mhan9@kennesaw.edu](mailto:mhan9@kennesaw.edu).

**QIXUAN HOU** is a master's degree student in analytics at the Georgia Institute of Technology, Atlanta. Her research interests are in the area of data science, with experience executing data-driven solutions. Hou received a B.S. in computer science and mathematics from the Georgia Institute of Technology. Contact her at [qhous6@gatech.edu](mailto:qhous6@gatech.edu).

**BAOZHONG YANG** organized the inaugural and second GSU-RFS FinTech Conferences and has served on the Program Committee of many conferences. His research interests include FinTech and investments. Yang received a Ph.D. from Stanford University, California, and the Massachusetts Institute of Technology, Cambridge. Contact him at [bzyang@gsu.edu](mailto:bzyang@gsu.edu).

framework," U.S. Patent 724 613 7B2, Jul 17, 2007.

7. R. S. Chen and C. M. Sun, "A collaborative continuous auditing model under service-oriented architecture environments," in *Proc. 6th WSEAS Int. Conf. E-ACTIVITIES*, 2007, pp. 47–52.
8. C. Wang, S. M. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in *Proc. IEEE INFOCOM*, Mar. 14–19, 2010, pp. 1–9. doi: 10.1109/INFOCOM.2010.5462173.
9. T. P. Pedersen, "Non-interactive and information-theoretic secure verifiable secret sharing," in *Proc. Advances Cryptology-Crypto'91*, 1991, pp. 129–140. doi:10.1007/3-540-46766-1-9.
10. F. Knirsch, A. Unterweger, and D. Engel, "Privacy-preserving blockchain-based electric vehicle charging with dynamic tariff decisions," *Comput. Sci. Res. Develop.*, vol. 33, no. 1–2, pp. 71–79, 2018. doi:10.1007/s00450-017-0348-5.
11. Y. Zhang, R. H. Deng, X. Liu, and D. Zheng, "Blockchain based efficient and robust fair payment for outsourcing services in cloud computing," *Inf. Sci.*, vol. 462, pp. 262–277, Sept. 2018. doi: 10.1016/j.ins.2018.06.018.
12. L. Xu et al., "Enabling the sharing economy: Privacy respecting contract based on public blockchain," in *Proc. ACM Workshop Blockchain, Cryptocurrencies and Contracts (BCC '17)*, 2017, pp. 15–21. doi:10.1145/3055518.3055527.
13. N. Szabo, "Formalizing and securing relationships on public networks," *First Monday*, vol. 2, no. 9, 1997. doi: 10.5210/fm.v2i9.548.

- .com/opinion/blockchain-is-already-changing-accounting
4. L. Coleman, "Big four giant PwC announces blockchain auditing service," *Yahoo! Finance*, Mar. 17, 2018. [Online]. Available: <https://finance.yahoo.com/news/big-four-giant-pwc-announces-124211024.html>

5. C. H. Wu, Y. E. Shao, B. Y. Ho, and T. Y. Chang, "On an agent-based architecture for collaborative continuous auditing," in *Proc. IEEE Int. Conf. Computer Supported Cooperative Work Design*, 2008, pp. 355–360. doi: 10.1109/CSCWD.2008.4536957.
6. S. Paulus, T. Schroer, and C. Buchholz, "Collaborative audit



# PharmaCrypt: Blockchain for Critical Pharmaceutical Industry to Counterfeit Drugs

**Neetesh Saxena**, Cardiff University

**Ieuan Thomas**, Bournemouth University

**Prosanta Gope**, University of Sheffield

**Pete Burnap**, Cardiff University

**Neeraj Kumar**, Thapar Institute of Engineering and Technology

*This research analyzes the impact of counterfeit drugs on the health-care supply chain industry and evaluates the solutions currently in place to reduce the number of counterfeits coming to the market. Feedback information obtained from industry professionals is used to build requirements for PharmaCrypt, a new blockchain-driven tool.*

**T**he presence of counterfeit medicines within the health-care industry is evident, with one in 10 medical products in developing countries being substandard or falsified (do not meet

the standards of safety, efficacy, and quality).<sup>1</sup> Falsified medicines can contain incorrect ingredients and doses or show no presence of the active ingredient. This means that there are millions of patients unaware that they are taking medicines that fail to work as prescribed. Not only do they fail to treat individuals, but some counterfeits can cause serious illness or even death. A modeling

Digital Object Identifier 10.1109/MC.2020.2989238  
Date of current version: 1 July 2020



exercise developed by the University of Edinburgh estimates that 72,000 to 169,000 children may be dying each year from pneumonia due to sub-standard and falsified antibiotics.<sup>1</sup> Counterfeits are a huge commercial drain for individuals and health-care systems and, in some cases, can lead to a further financial burden on the health-care system if the patient consequently requires treatment.

### PROBLEM STATEMENT

The health-care industry is rife with counterfeit drugs (which infringe upon the intellectual property rights) that penetrate the industry's supply chain.<sup>2</sup> This is due to a complex supply chain, compounded by a lack of visibility of the products' end-to-end journey. The effects of falsified and counterfeit drugs have the potential to cause devastating consequences.

The complexity of the drug distribution supply chain makes it difficult to prevent counterfeits infiltrating the industry. There are numerous companies involved in the supply chain where drugs change ownership between manufacturers to distributors, repackagers, and wholesalers before reaching the patient. There is little or no visibility between the parties involved in the supply chain to track the authenticity of the drugs. This causes a level of uncertainty for patients and dispensaries concerning the authenticity of the products sold at the end of the chain.

There are currently several solutions to this problem, but as the sophistication of counterfeit products and packaging rapidly improves, they have flaws and limitations. Some solutions endeavor to trace transactions of the products as they move through the supply chain and change ownership, although there is still a central

organization present that is at risk of being compromised, whereby documents can easily be falsified. Also, a central system is prone to a single point of failure. Solutions like our proposal could potentially be adapted to include the antitampering and distributed database capacities of blockchain.

### PROPOSED SOLUTION AND CONTRIBUTIONS

The proposed solution presents a blockchain-driven tool that can be used to record and time-stamp the transfer of goods at each point in the pharmaceutical supply chain. As the drug travels through the supply chain, every transaction of goods will be noted and time-stamped by scanning the bar code. The ledger will be used to ensure the security and safety of the product. The threefold contributions of this article are as follows:

- › analysis of the problem and a demonstration of whether using blockchain could be a better solution to the supply chain of drugs than the existing solutions
- › design and creation of an application tool that can be used to record the origin and contents of the drugs manufactured and time-stamp the transfer of goods
- › provide recommendations based on the tool's functioning as to whether (and how) using blockchain technology is the best way to solve this problem.

### BACKGROUND STUDY

This section discusses the problem domain in greater depth and highlights existing solutions and evaluates their limitations. We will also deliberate as to why a blockchain solution

could be an improved idea compared to the current solutions.

### Drug supply chain in the pharmaceutical industry

The pharmaceutical supply chain is the means by which prescription medicines are delivered to patients.<sup>3</sup> Ingredients for medicines are normally sourced from a variety of places before reaching the final formula. Once the final formula is achieved, the drug can be distributed. During the supply chain lifecycle, the drug will transfer among many different entities, specifically between the manufacturer and the patient. Every transaction offers an opportunity for counterfeit or falsified products to penetrate the supply chain and the industry. Figure 1 shows a typical supply chain scenario in the pharmaceutical industry.

**Manufacturers, wholesalers, and pharmacies.** The manufacturer's role within the supply chain is to ensure the readiness of its inventory of drugs so they can be distributed to wholesalers. Manufacturers receive orders from distributors/wholesalers and then ship the products to the distributor's warehouses where they are put away in storage. Distributors provide manufacturers with inventory data reports to maintain transparency throughout the process.

The role of wholesalers is to make the process of purchasing pharmaceutical drugs a simpler and more efficient process. Wholesalers connect and deliver to thousands of pharmacies and dispensers. This saves manufacturers efforts of dispatching drugs to pharmacies individually, because instead they can send large batches of medications to a relatively smaller number of wholesalers. Once the product is in the hands of the wholesaler, it

can provide a range of services, including drug distribution, electronic order services, and repackaging.

The final entities in the supply chain are pharmacies and hospitals. Pharmacies account for approximately 75% of the prescription drug market, whereas such nonretail providers as hospitals comprise the remaining 25%.<sup>4</sup> Pharmacies and hospitals purchase products from wholesalers, and then they are sold to the final patient.

**The wholesaler problem.** Primary wholesalers have direct distribution contracts with the manufacturers they purchase from, whereas secondary wholesalers purchase products from a range of other parties. Shown by the arrows between the two wholesalers in Figure 1, it may not always be obvious whether a company is a secondary or a primary wholesaler. For example, a primary wholesaler may not only purchase products directly from the manufacturer; it may also purchase from secondary wholesalers depending on the demand for certain medicines. Buying and selling between wholesalers are common within the industry, and products move between a variety of different companies and can be repeatedly repackaged by each wholesaler before reaching the patient.

In a process called *sating*, counterfeit drugs can be merged and confused with legitimate products at the wholesalers. This can be unknowingly caused if, for example, a wholesaler purchases from a secondary wholesaler company that has accidentally purchased counterfeit goods. During the repackaging process at the wholesaler, the counterfeit drugs may be given genuine labels. Manufacturers initially deliver medicines in fraud-protection packaging. This can be removed during the

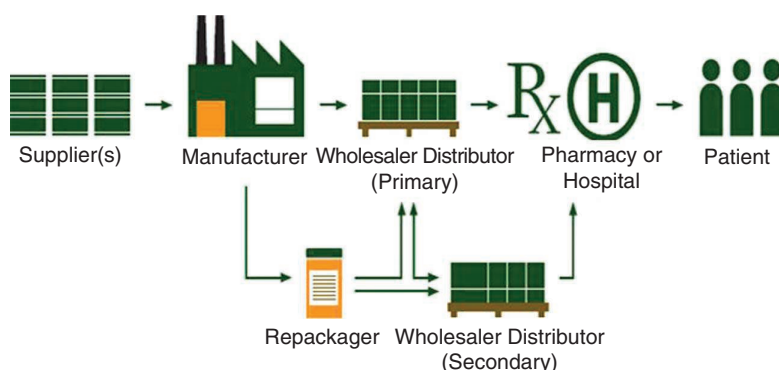


FIGURE 1. The pharmaceutical supply chain.

repackaging phase, and batch numbers may be reprinted.

**Drug diversion.** Drug diversion occurs when drugs that have been authorized to be sold in one country are sold in another. Criminals take advantage of segments in the supply chain where products leave a documented chain of custody and they can implant falsified goods. Markets that trade in diverted drugs usually have little oversight from authorities and are known as *gray markets*.

### Existing solutions

This section briefly discusses the current solutions that are in place.

**Packaging.** In an attempt to reduce the impact of counterfeit drugs, several pharmaceutical companies have adopted a more sophisticated packaging approach. One of these approaches is the use of holographic technologies. The concept is that a patient will know if the product is legitimate when he or she sees that the packaging contains a hologram. A major advantage of this type of packaging is that it can be applied to every individual item. This type of packaging can be costly to implement depending on the

complexity of the hologram. Holograms can also eventually be cloned by counterfeit companies, making the original secure packaging ineffective. Another disadvantage to this solution is that it does not offer companies intelligence for when a counterfeit product penetrates the supply chain.

**Mass serialization.** Mass serialization is a technology used to identify and track objects and individuals using radio-frequency (RF) waves. Manufacturers can use RF identification (RFID) coding to allocate packages with unique identifiers. As the product makes its way through the supply chain, the product's information is captured by a chip reader. The chips can be disguised within large batches of products to avoid tampering. However, RFID is costly to implement because the RFID tags themselves are expensive. There are many types of systems that include varied readers and tags, creating compatibility issues.<sup>5</sup> Another concern with the technology is that it has the potential to be hacked and information on the tags can potentially be altered.<sup>6</sup>

**Mass encryption technology.** Software-based mass encryption technology

TABLE 1. A comparison of existing solutions.

Solution	Pros	Cons
Packaging	It is easy for patients to determine whether the drug is legitimate if they see a hologram on the packaging.	It is expensive and can be cloned; the origin of fraudulent products cannot be located.
Mass serialization	It has the ability to track and trace; chips can be disguised in large batches to avoid tampering.	It is costly, has potential to be hacked, and has compatibility issues; chips have the potential to be tampered with.
Mass encryption technology	Each batch is given a unique code.	It is complicated to implement.
FMD	An EU-wide directive sets the standards for all manufacturers.	Packaging can be forged and overseen by a centralized authority that could be liable to attack.

can be used in the pharmaceutical industry to fight against counterfeit drugs. The same software is required to decrypt the digital code.<sup>7</sup> This technology requires a large database server to store the data.

**Falsified Medicines Directive: Safety features.** The Falsified Medicines Directive (FMD) is a European Union (EU) directive that aims to ensure that medicines in the EU are safe and a trade of them is properly controlled by including a unique identifier and an antitampering device on the packaging.<sup>8</sup> As the product goes through the supply chain, at various points it is mandatory that the bar code is scanned again. This aims to certify the authenticity and trustworthiness of the medicine supplied. The unique identifier on the packaging must encompass a product code determining the medicine name, common name, pharmaceutical form (strength, pack size, and pack types), serial number, batch number, and expiry date. Manufacturers were required to comply with this EU directive beginning on 9 February 2019. Table 1 displays the pros and cons of the current solutions discussed.

WHY BLOCKCHAIN?

This section explains why we believe that blockchain is the suitable solution.

Blockchain—Technical feasibility

Blockchain is a digital technology model that can be used to store data. It consists of a chain of blocks containing transaction information.<sup>9</sup> It is a decentralized system where data can be shared across a network in an encrypted fashion.

Before a transaction can be added to a blockchain, it needs to be verified by the network nodes using a majority consensus protocol, where nodes on the network agree that the transaction is legitimate. Any transactions that have been recorded cannot be altered or erased, and the full transactional history can be viewed at any time. Each block in the chain contains data: the hash of the block, hash value of the previous block, and nonce of the existing block. Hashing is used to make integrity-protected blocks together to create the secure chain. Every data block in the blockchain is given a unique digital signature that directly corresponds to the data in its block (the hash). If the data in the block are

changed, the digital signature of the block will also subsequently change.

A block registers transactions as they occur, and the blockchain increases in size periodically as new transactions execute. Once the block is filled, it is allocated a digital signature that directly corresponds to the string of data in that block (hash). The first block in the chain is known as the *genesis block* and does not point to any previous blocks. To link another block of transaction data, the signature in the first block is added to the data of the following block. The digital signature of the second block is now partially dependent on the signature of the block before it, as it is included in the data of the block. This process is repeated every time new transactions occur to create the chain.

Blockchain platforms

We discuss three main platforms that were considered: the Ethereum app platform, Amazon Web Services (AWS), and Oracle blockchain.

**Ethereum app platform.** Ethereum is a publicly distributed blockchain network that provides users with the appropriate environment to



deploy decentralized applications. The platform runs the smart contracts that have been set by the application developer. The Ethereum network is made up of a series of distributed nodes and Ethereum wallets. The distributed network of nodes is established when computers or miners join the network. The network does not hold any permissions to join, as any node with enough computing power is able to join the network.

**AWS.** AWS provides cloud computing platforms. Amazon offers blockchain templates as a part of its platform, which provides users a simple way to build blockchain applications for businesses. AWS provides the ledger database behind the application that eradicates the need for the application owner to develop the complex blockchain network. The service offers two types of use cases: to track and verify transactions with centralized ownership and to execute transactions and contracts with decentralized ownership.<sup>10</sup> Using the AWS blockchain template, an Ethereum blockchain network on a cluster made up of multiple instances with an application load balancer (ALB) can be created. We have used the AWS platform because its service makes it easy to set up, deploy, and manage scalable blockchain networks, which eliminates the need to rely on other expensive implementations.

**Oracle blockchain.** Oracle is extremely similar to AWS because it is a blockchain-as-a-service provider. It allows businesses to deploy applications over an immutable electronic distributed ledger database.

### Blockchain security

Blockchain not only allows users to integrate with suppliers, customers, regulatory agencies, and stakeholders

but also provides a high degree of accuracy.<sup>6</sup> It also offers a higher level of security compared to the existing solutions.

#### Immutability and consensus.

The immutable characteristic of blockchain is one of the main reasons companies are starting to implement the technology. If a block is altered, it will unchain itself from the consecutive blocks. For an altered block to be accepted on the blockchain, it needs to be chained to the rest of the blocks. All of the nodes in the network work together to create a consensus about which blocks are valid and which are not. Users in the blockchain will be notified that data have been altered and will deny the change. The blockchain will then be returned to a previous state where all blocks are still chained together.

**Private keys.** Participating nodes in the network are assigned their own private keys that are linked to transactions they make. The private key is used to create a digital signature and sign each transaction. Each node in the network is allocated a private key, which grants ownership to its data entry.<sup>11</sup>

**Decentralization.** Rather than relying on a single database to secure transactions with users, blockchain is completely decentralized. This means there is no single point of failure, there are multiple copies of the same transactions, and a hacker would need to change all copies and break the consensus protocol before anything could be altered.

#### Existing solution case study: Cisco supply chain management

Cisco loses out on more than US\$500 million dollars of revenue a year due to counterfeit products (similar to the United Kingdom's losses of £218 million every year from counterfeit wine

and spirits<sup>12</sup>); clearly it is not just a problem faced by developing countries. As a result, Cisco is currently working on a blockchain solution designed to combat against counterfeit products on its own supply chain. Although Cisco works in a different industry than the one discussed in this article, its application of the technology is extremely similar. A few other enterprise blockchain use cases include<sup>13</sup> supply chain management (IBM Food Trust), protecting digital identity (Civic's Secure Identity Platform), smarter predictive analyses (Endor), and health-care medical history and records (Medicalchain).

There are some limitations with the technology, and we must know how these drawbacks might be overcome. As the participation of each organization in the supply chain requires complex infrastructure to be able to run a single node, one of the main issues with blockchain is the cost of this infrastructure. It is hard for customers and suppliers to justify the cost. A possible resolution for smaller companies could be to provide a cloud-based solution, but it is not quite clear how well this scales up. There is currently much research being carried out to work out ways of reducing the costs and monetizing the process of blockchain.

Another limitation of blockchain arises when considering the consensus mechanism used. There are multiple consensus protocols available. In a public blockchain, it is possible to specify a single consensus mechanism used, although in a private enterprise solution, it is not possible to make an application as rigid. There is ongoing research on how to make the consensus as quick as possible and pluggable so that suppliers can appoint the consensus they wish to use. For a blockchain enterprise application to

### INTERVIEW QUESTIONS AND ANSWERS

We interviewed 30 people who are directly or indirectly working in the pharmaceutical industry.

Q1: Are you aware of the counterfeit drugs problem in the pharmaceutical industry?

A1: 1) Yes. ×19

2) Yes, but I believe it to be more of a problem in developing countries than in the UK. ×4

3) Yes I am aware, but I have never experienced it myself. ×5

4) Yes. As a chief pharmacist of an NHS [National Health Service] trust and a responsible person on a wholesale dealer's license from the MHRA [the Medicines and Healthcare Products Regulatory Agency], I am acutely aware of the potential for falsified medicines entering the supply chain. ×1

Q2: Do you know if there are any products or systems in place that are used to track and trace a drug through its supply chain before it gets to the dispensary? If yes, could you explain what?

A2: 1) I don't know. ×4

2) The FMD is currently being implemented. ×9

3) EU directive. ×2

4) We use registered wholesalers; however, the only system currently being implemented is FMD scanning. ×10

5) Scanning the products—bar codes, QR codes. ×4

6) Up until recently, some products have had 2D bar codes, holograms, and tamper-evident packaging to reduce falsification, although the sophistication of counterfeiters now is such that even these can be replicated. The main intervention now is the introduction of the Falsified Medicines Directive, which requires licensed medicines to have a 3D bar code, a unique identification number traceable to individual packs, and tamper-evident packaging. Each individual pack is tracked via a Europe-wide repository. ×1

Q3: Are there any systems/methods in place you use personally that help to ensure the authenticity of the medicines supplied to customers? If yes, could you explain what?

A3: 1) Medicines are scanned in on arrival, but this is more for stock check purposes than authenticity. MHRA supply regularly alerts us if there are any concerns regarding medicines, and these are relayed to all pharmacies and dispensaries with relevant batch numbers. ×2

2) No. ×6

3) Ensuring everything we order is done through our trusted suppliers we use. ×2

4) FMD scanners. ×15

5) Scanner but limited due to possible human error. ×4

6) Only those already mandated.

be successful, the consensus times need ideally to be minutes or seconds. When choosing a consensus mechanism, it is important that the protocol is Byzantine fault tolerant.

#### PROPOSED SOLUTION: REQUIREMENT AND ANALYSIS

During the analysis stage, professionals working on or researching blockchain as well as those in the pharmaceutical field were contacted and engaged

in collaboration. The purpose was to determine the use cases of blockchain currently in practice in the industry, people's conceptions of the technology, and how staff in the pharmaceutical industry might cope with or react to a new supply chain management system.

#### Pharmaceutical interviews/ discussions

The participants chosen were individuals who work in the health-care

industry either in a dispensary/pharmacy or as a pharmacist.

#### Pharmaceutical feedback

We interviewed 30 people who are directly or indirectly working in the pharmaceutical industry. Roughly 60% said they were aware of the counterfeit drugs problem: "Medicines are scanned on arrival, but this is more for stock check purposes than authenticity. MHRA [the Medicines and Healthcare

Q4: What information about a medicine would you suggest needed to be logged to ensure its authenticity?

A4: 1) Batch number, expiry date, manufacturer. ×7

2) If you look at what the Falsified Medicines [Directive] safety features require, it might give you a good idea on what to include. Off the top of my head, I believe it is required that manufacturers provide the name, serial number, expiry date, strength, and batch number but there may be more. ×2

3) Special packaging. ×3

4) Name, batch number, wholesaler. ×9

5) Ingredient constituent and manufacturer who has approved it.

6) Product, batch number, expiry, product license number, manufacturer. ×8

Q5: Would you consider using an application to ensure the authenticity of medicines if it meant scanning the bar code of each drug sold over the counter?

A5: 1) Yes. ×17

2) Yes, but it's a hassle. ×7

3) Possibly, depending on the efficiency of the system. ×3

4) We already scan products so yes. ×2

5) This is effectively what FMD requires for prescription medicines. The same principle for over-the-counter medicines would probably work OK where the process can be combined with another (for example, scanning at point of sale). ×1

Q6: Are you aware of blockchain or cryptocurrency technology, that is, bitcoin?

A6: 1) Yes. ×7

2) No. ×4

3) Have heard of it but never used it. ×19

Q7: Would you put your trust in a blockchain-driven application that was designed to track and trace a medicine as it makes its way through the pharmaceutical supply chain? (Blockchain is the technology behind cryptocurrencies like bitcoin.)

A7: 1) Yes. ×12

2) No. ×8

3) Unsure. ×9

4) Would consider using it but would need robust evidence and assurance before trusting it completely. ×1

Products Regulatory Agency] regularly alerts us if there are any concerns regarding medicines, and these are relayed to all pharmacies and dispensaries with relevant batch numbers. If you look at what the [FMD] safety features require, it might give you a good idea on what to include. It is required that manufacturers provide the name, serial number, expiry date, strength, and batch number but there may be more." In responding to whether they would trust a

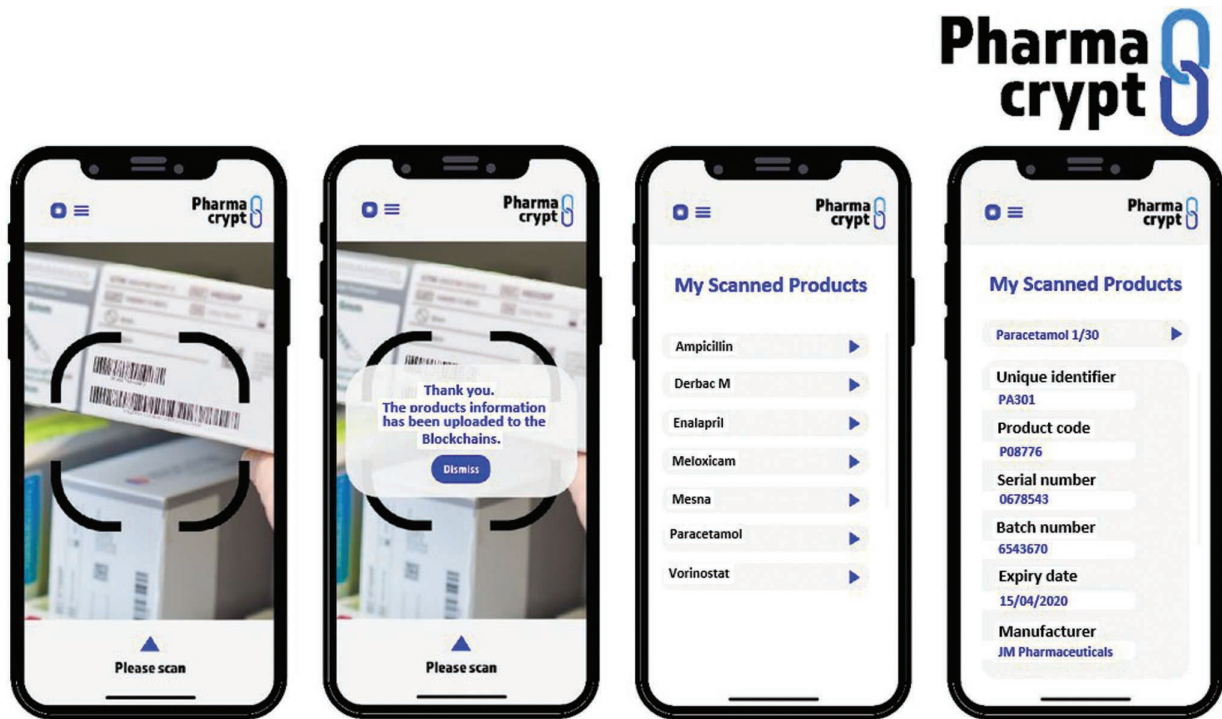
blockchain-driven application designed to track and trace a medicine as it makes its way through the pharmaceutical supply chain, most people were interested. The information gathered indicated that the most important material pharmacists perceived as required by an application (tool) to ensure the trustworthiness of a product would be batch number, name, expiry date, and manufacturer. The majority of people interviewed suggested bar code scanning is

already used in their day-to-day job and implied that they would consider using an application (tool) that requires the use of bar code scanning. The original data collected from the focus group and interviews can be found in "Interview Questions and Answers."

## SETUP AND THE PHARMACRYPT TOOL

This section explains the experimental setup for the development of the





**FIGURE 2.** The proposed PharmaCrypt application tool user interface.

proposed PharmaCrypt application tool using Ethereum blockchain, which is created using AWS. The tool interface can be seen in Figure 2. The network is used to create a smart contract where products can be created and transferred between accounts.

### PharmaCrypt features

The developed prototype of the tool has the following features:

- › **Bar code scan:** Handheld smart-phone devices are able to scan bar codes and upload the information to the blockchain.
- › **Asset creation:** The application is able to create new assets when products first enter the supply chain. A product's information

is uploaded to the block-chain and assigned a unique identifier number.

- › **Transfer of asset:** When a product is moved on to the next supplier or entity in the supply chain, the application tool records the transaction.
- › **View scanned products:** A user is able to view all user-scanned products.
- › **Performance requirements:** The application tool scans bar codes instantly with no lags or glitches. Consensus should be reached in under a few seconds.
- › **Security requirements:** 1) Separate accounts for each user; 2) users are enrolled with

business network accounts; 3) each password is at least eight characters long and composed of at least one uppercase letter, a number, and one special character; and 4) each user is operated using the least set of privileges required to do his or her job.

### Blockchain implementation

AWS was chosen to create the prototype of the proposed tool. The overall software is hosted on a computer-based system, configured with an i7 processor, 500-gigabyte hard disk drive, 4-gigabyte random-access memory (RAM), and Windows 10 operating system, and the required data can be fetched to a mobile application. The

AWS blockchain for Ethereum creates a private blockchain network on the AWS CloudFormation. The final network is made up of the following entities: two Ethereum clients, one miner running on Amazon Elastic Compute Cloud (EC2) instances in an Amazon EC3 cluster, on-demand EC2 instances, and an internal ALB. The entire process of the proposed solution and building the PharmaCrypt tool is as follows:

1. *PharmaCrypt tool interface:* As the functions of the application tool run on smartphone devices are relatively limited, the user interface, as shown in Figure 2, will largely be the same for each company. Here, the user will be able to scan the bar code of the product. Once the bar code has been scanned, if the transactional data are deemed legitimate by the network, they will be automatically uploaded to the blockchain. Users of this interface will be able to view only the transactions they have scanned themselves. If the transaction is deemed illegitimate by the blockchain network, an error message will take over the screen. The error message will trigger a notification sent to the main computer-based interface controlled by senior personnel. The supply chain management team can then investigate this product further. Figure 3 explains the information flow of the proposed solution.
2. *Key-pair generation:* AWS uses public-key cryptography to secure the login information of the instances in the network. As shown in Figure 4(a),

we created a key pair for the blockchain Ethereum network, which is used to sign every transaction over the network. The key pair must be created in the same region in which you wish to launch the instance. The key pair will download, and the file name is the name you specified with a .pem extension.

3. *Subnets, security groups, and rules:* The Amazon Virtual Private Cloud (VPC) is used to define the virtual network where resources will be launched. An ALB is created requiring two public subnets to be configured located in two separate availability zones. A private subnet is also necessary for the container instances. The availability zone should be located in the same zone as the ALB.

AWS security groups control the inbound and outbound traffic to your resources. We specify two security groups: one for controlling the traffic between the EC2 instances in the cluster and the other for controlling the traffic between the ALB, EC2 instances, and bastion host. Thereafter, we have applied the following incoming rules to these groups, as shown in Figure 4(b):

- Allow all traffic from the ALB security group permitting the ALB to broadcast with itself and the bastion host.
- Allow all traffic from the EC2 security group permitting instances in the security group to broadcast to the ALB and the bastion host.
- Allow Secure Shell (SSH) traffic from the Internet Protocol

address that permits traffic from the computer to the bastion host.

The following outbound rules also need to be applied on the same security group:

- Allow all traffic from the EC2 security group that permits outbound traffic from the ALB and the bastion host to the instance.
  - Allow all traffic from the ALB security group that permits the ALB to communicate with the bastion host and itself.
4. *Identity and access management and bastion host:* We created a role for AWS service selecting Elastic Container Service (ECS) for the service and ECS for the use case. Make a note of the role—Amazon Resource Name, as it will be needed later. The bastion host is an instance used to connect to the web interfaces and other instances in the network. To do so, the bastion host forwards SSH traffic from trusted clients that are outside the VPC.
  5. *CloudFormation stack:* Now the tool has been configured, and the Ethereum network can be created. To do so, an AWS CloudFormation stack needs to be set up. The AWS CloudFormation stack establishes an Amazon EC3 cluster of EC3 instances. Launching this stack creates some nested stacks where we are able to connect to the network resources using the bastion

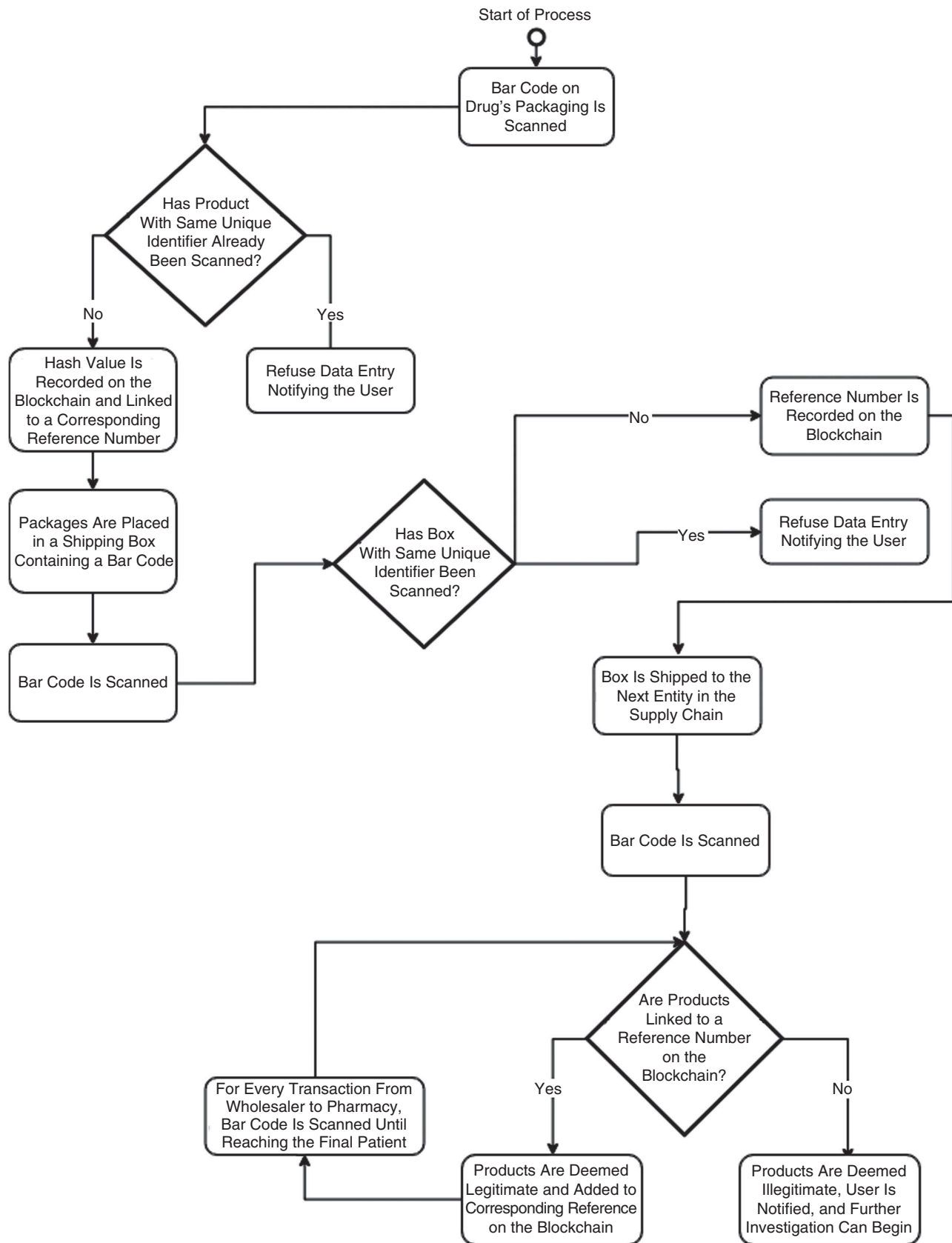


FIGURE 3. The proposed solution information flow.



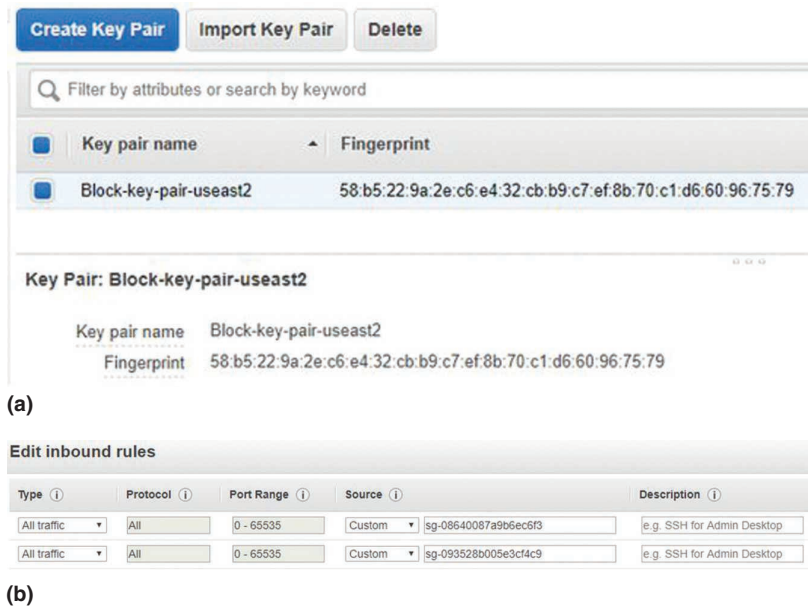
host. In the dashboard, their progress can be observed by selecting Stacks. When the stacks have finished creating, the Output tab displays Ethereum uniform resource locators (URLs) we can connect to where the EthStats (shows the time since something was mined), EthExplorer (block-chain explorer), and EthJson-RPC [a stateless, lightweight JavaScript Object Notation (JSON) Remote Procedure Call (RPC)] are displayed.

6. *Connect SSH port, authenticate, and set up a proxy:* Now, to connect to the bastion host, an SSH port forwarding connection is established using PuTTY. The key pair needs to be converted to a .ppk format because PuTTY does not support the default .pem format. We have used the RSA algorithm for the key generation.

We set the following configurations: select Connection, SSH, Tunnels; add 9001 as the source port, and leave destination as default (blank). Thereafter, use Open to authenticate the bastion host. We then configured a proxy (FoxyProxy for Chrome browser) on port 9001 so that the forwarded port can be used to connect to the Ethereum URLs.

The EthStats URL displays the status of the Ethereum network. The EthExplorer URL where transactions have been made on the network is shown in Figure 5.

7. *Smart contract, genesis block creation, and mining:* Now, we need to create smart contracts



**FIGURE 4.** The (a) key-pair creation for signing the transaction data and (b) security group control: inbound rules for the traffic.

and run them (with admin permission) on the blockchain network, as shown in Figure 6. To do this, we connect over to the Windows bastion host using Remote Desktop Protocol with a decrypted password (using .pem private key). Ethereum Wallet (allows management of bitcoin, Ethereum, XRP, and more than 300 coins and tokens) and geth (a command line interface for running a full Ethereum node implementation) are not designed to securely connect to the remote nodes in the network using RPC. To set up a secure connection, we run a local geth node that joins the network.

The genesis block is the first block in the blockchain

network. The genesis block must be compatible with the private blockchain network that has been created. Creating a genesis block allows you to sync the node with the network. To do so, define static node mapping in a JSON file using the information available in the Amazon DynamoDB table.

Now, the geth client needs to be initialized to use the genesis block you constructed. Thereafter, use the Ethereum Wallet app to store the keys, contracts, tokens, and ether. We have used mining with two threats for this demonstration work. Now we have Ether, so we deploy the first smart contract product tracker (Ethereum Wallet application → Contracts → Deploy New



FIGURE 5. The (a) Ethereum blockchain network with block details (EthExplorer URL) and (b) Ethereum blockchain real-time network status.

Contract). The smart contract uses the Solidity coding language to create (using the Remix tool) and transfer assets on the Ethereum blockchain network.

Discussion

To demonstrate the working of this tool, we used permissioned blockchain, which is in fact more scalable and faster but works toward centralized controls among a group of users (who were involved in this article). However, this work can be easily extended to permissionless blockchain, when required, so that any registered user can validate transaction information, and this will be

tested with trials. We have tested this tool with 50 users making transactions on different items to demonstrate that it is useful for small and medium-size pharmaceutical applications, and we will further extend its capabilities (by testing) for the applications with a wide variety of drugs and a large number of users involved in the system. We have performed mining on the cloud without an application-specified integrated circuit miner. It does not yield any profit, but the primary purpose is to demonstrate the working of this prototype tool, which at present will be used privately with a limited number of drugs and users involved in the system.

Comparison

Compared to Drugledger by Huang et al.,<sup>11</sup> the proposed consumer-oriented application tool provides more controllability, a more user-friendly interface, added security through groups, and private network virtualization. The proposed PharmaCrypt does not require a certificate service provider, antiattack service provider, and query service provider, which are the requirements for Drugledger. In other words, the proposed tool generates less overhead compared to Drugledger. A detailed comparison between PharmaCrypt and Drugledger is presented in Table 2 and Figure 7.

Apart from this, a smart contract using smart storage containers proposed by Hinckeldeyn and Jochen<sup>14</sup> is based on a multisignature wallet of three parties to process the payment and arbitrate disagreements. The application requires extra hardware and protocol implementation, which is time-consuming as well as inefficient. Similar to the approach by Mondal et al.,<sup>6</sup> the proposed tool is able to defeat tampering, spamming, physical layer attack, and preferential treatment. However, the approach by Mondal et al.<sup>6</sup> requires new sensors, their set up, and significantly large storage for the algorithms to run, whereas PharmaCrypt does not rely on extra sensors; rather it uses existing technology of scanning the bar codes. The only drawback of the proposed tool is that it may be affected by the service provided by AWS, because it is based on the AWS blockchain.

A performance comparison between the PharmaCrypt and Drugledger is shown in Figure 7. The system used is 64-bit Windows 10 with core Intel i5 2.60 GHz and 4-gigabyte RAM with Java.

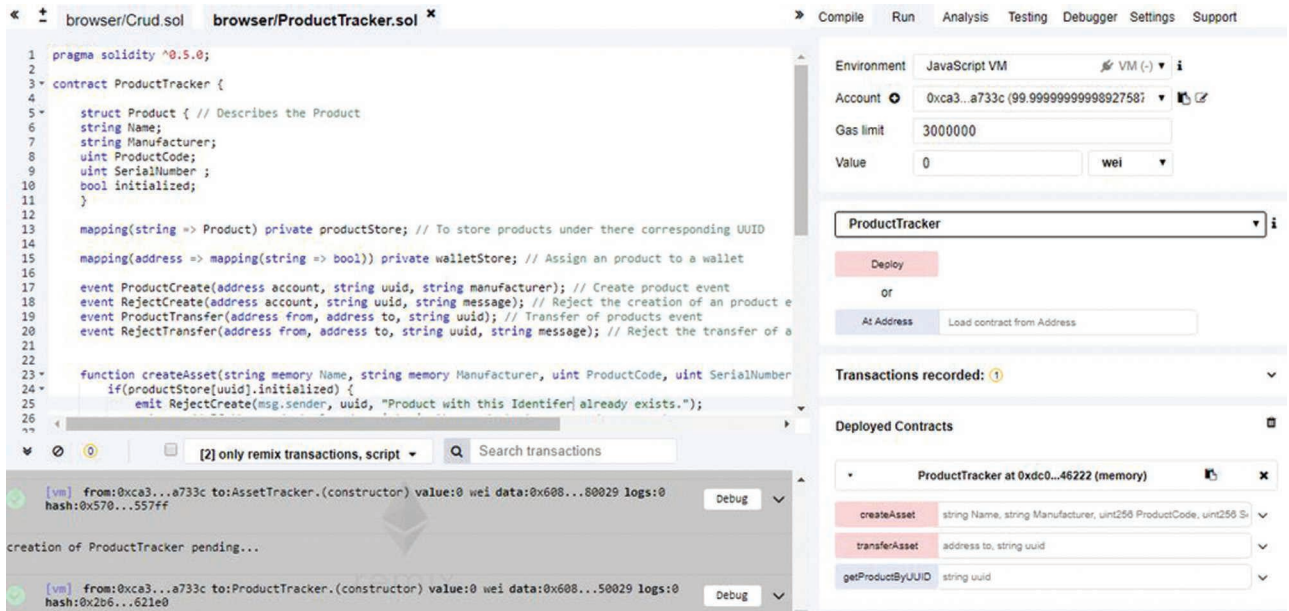


FIGURE 6. A smart contract in remix.

For generating random strings, we have considered `UUID.randomUUID().toString()` and used `System.currentTimeMillis()` for calculating time stamps and execution times. `SHA256()` takes 20 ms for generating hash code for each block. Further, it takes 1 ms each to create the block lifetime, Merkle root, time stamp, and version. Due to the insufficient implementation details available in Drugledger<sup>11</sup> [the details are not provided, so we created simple functions for `SynchronizeUTXO()`—20ms, `ReadDrugPackage()`—10 ms, `GetUTXO()`—20 ms, `ValidQuery()`—20 ms, `CreateTX()`—20 ms, `Gossip()`—10 ms, and `IsCorrelated()`—10 ms], we have assumed a similar type of parameters and size as PharmaCrypt. Figure 7(a) demonstrates the average block time when the number of blocks is 10, 50, 100, and 200. Figure 7(b) reflects comparative instances of average

block time when the number of blocks is 10, 20, 30, 40, and 50 for PharmaCrypt and Drugledger. Overall, it is clear that PharmaCrypt outperforms

near-field communication and blockchain technologies, and 2) a decentralized consensus protocol. However, it is not clear whether this protocol can be used

**WE HAVE TESTED THIS TOOL WITH 50 USERS MAKING TRANSACTIONS ON DIFFERENT ITEMS TO DEMONSTRATE THAT IT IS USEFUL FOR SMALL AND MEDIUM-SIZE PHARMACEUTICAL APPLICATIONS.**

Drugledger and is better suited for such an application.

Alzahrani and Bulusu<sup>15</sup> proposed 1) Block Supply, a decentralized anticounterfeiting supply chain that is based on

for such blockchain-related applications as PharmaCrypt. Wang et al.<sup>16</sup> combined the emerging blockchain technology with parallel health-care systems for comprehensive health-care data



TABLE 2. A comparison of blockchain solutions.

Solution	Drugledger <sup>11</sup>	PharmaCrypt
Basic requirements included	More focus on packaging and repackaging (overall less efficient)	More focus on rapid scanning of the product (bar code scanning), asset creation, and transfer (overall more efficient)
Overhead	High, due to the maintenance of certificates, per-transaction user weight computation, and repackaging	Low, none required
Technology	Platform dependent, C++ in Ubuntu 16.04 long-term support	Platform independent with AWS
Extra requirements for security support	Requires certificate of service provider	Free from such requirement
Performance (efficiency)	Not specifically discussed but much slower	Improved using bar code scanning, average block time 2.11 s, page latency 2 ms, and consensus available in a few seconds
Security key and hash storage	There are issues with storing the public key and hash codes	There are no such issues because the AWS storage takes care of it

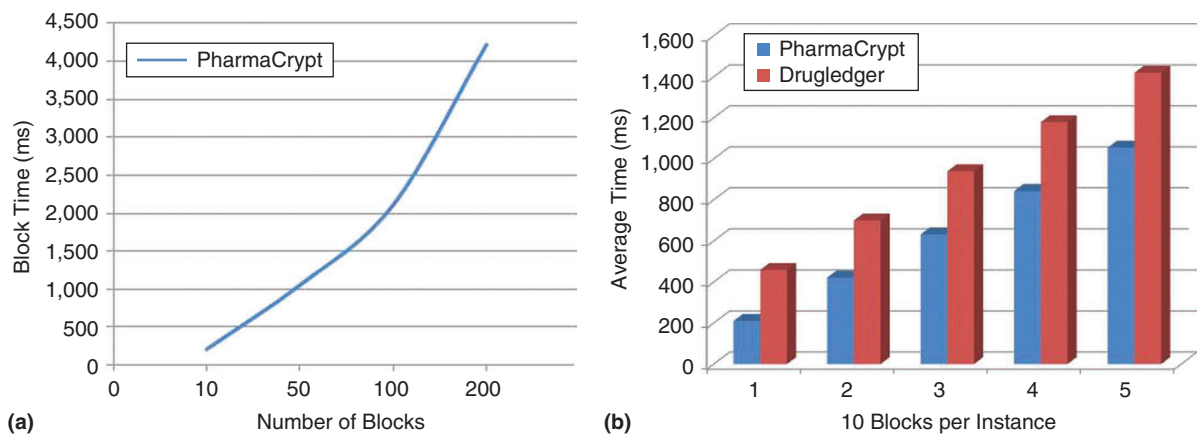


FIGURE 7. The details on average block time. (a) PharmaCrypt average block time where the number of blocks ranges from 10 to 200. (b) PharmaCrypt versus Drugledger block time where the number of blocks ranges from 10 to 50.

sharing, medical records review, and care auditability. The implementation aspect is not discussed in detail. Jamil et al.<sup>17</sup> proposed a novel drug supply chain management using Hyperledger Fabric based on blockchain technology to handle secure drug supply chain records. The

open source framework Hyperledger Fabric is used, but it is not clear whether it supports existing health-care systems and their services. We will further extend this research and see if any of this work can be extended and integrated with PharmaCrypt.

This work analyzed the counterfeit drug problem and existing solutions and evaluating their effectiveness. The inputs from relevant industry professionals working in both the pharmaceuticals industry and blockchain technology are

considered, which has actually helped to scope the requirements for the proposed application tool. In our primary research, 100% of the pharmacists interviewed were aware of the counterfeit drug problem, underlining just how widespread and severe the issue is within the health-care industry. Work is being done to fight the issue; however, the current solutions have a number of problems and limitations.

Further research is required to look at how we can achieve the shortest amount of time for a transaction to gain consensus. When using an application such as the one described, it is important that this time is as low as possible; otherwise, it will not be efficient for suppliers to use. Another need for further research would be to look at how it might be possible to lower the cost of implementation and understand how the solutions may drive down other supply chain operational costs in the pharmaceutical industry so that the technology is commercially viable for larger-enterprise solutions. Furthermore, the supply chain management system could be linked to a wider solution. There is currently work being undertaken to develop an electronic patient record system that can be used to store patient records on a blockchain. This system could potentially be combined with a supply chain solution whereby records in the blockchain could contain both patient treatment records and prescription history.

The EU FMD and RFID technology are currently the most effective in addressing the problem. The blockchain solution would be able to incorporate the compliance regulations so that the tool logs and tracks the information needed to comply with the directive. The way in which the blockchain tool is used

can mimic the RFID mass serialization process, and scanning of products can be carried out at the same points in the supply chain. This should enable a smooth transition to the new technology. Using PharmaCrypt, the proposed solution, means that both patients and dispensaries will be made certain of the provenance of the drug. The developed tool is relatively simple, meaning staff should not need extensive training due to existing product-scanning experience in dispensaries.

Blockchain also has its limitations, including a scalability issue. At this stage, it would be difficult to deploy a blockchain solution to all parties involved in the supply chain. Large-scale deployments across multiple customers would require much more rigorous testing to ensure success.

We believe that the proposed PharmaCrypt application tool has the capacity to be a successful working service and can be used as the basis for further research and development. ■

## REFERENCES

1. "1 in 10 medical products in developing countries is substandard or falsified," World Health Organization, Geneva, 2017. [Online]. Available: <https://www.who.int/news-room/detail/28-11-2017-1-in-10-medical-products-in-developing-countries-is-substandard-or-falsified>
2. T. Bocek, B. B. Rodrigues, T. Strasser, and B. Stiller, "Blockchains everywhere: A use-case of blockchains in the pharma supply-chain," in *Proc. IFIP/IEEE Symp. Integrated Network and Service Management (IM)*, Lisbon, Portugal, May 8–12, 2017, pp. 772–777. doi: 10.23919/INM.2017.7987376.
3. D. Yue, X. Wu, and J. Bai, "RFID application framework for pharmaceutical supply chain," in *Proc. IEEE Int. Conf. Service Operations and Logistics, and Informatics*, Beijing, 2008, pp. 1125–1130. doi: 10.1109/SOLI.2008.4686568.
4. M. Schapranow, C. Faehnrich, A. Zeier and H. Plattner, "Simulation of RFID-aided supply chains: Case study of the pharmaceutical supply chain," in *Proc. Int. Conf. Computational Intelligence, Modelling and Simulation*, Langkawi, Malaysia, 2011, pp. 340–345. doi: 10.1109/CIMSim.2011.68.
5. P. Behner, M.-L. Hecht, and F. Wahl, "Fighting counterfeit pharmaceuticals." PWC, 2017. [Online]. Available: <https://www.strategyand.pwc.com/gx/en/insights/2017/fighting-counterfeit-pharmaceuticals/fighting-counterfeit-pharmaceuticals.pdf>
6. S. Mondal, K. P. Wijewardena, S. Karuppuswami, N. Kriti, D. Kumar, and P. Chahal, "Blockchain inspired RFID-based information architecture for food supply chain," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5803–5813, June 2019. doi: 10.1109/JIOT.2019.2907658.
7. S. Dechand, A. Naiakshina, A. Danilova, and M. Smith, "In encryption we don't trust: The effect of end-to-end encryption to the masses on user perception," in *Proc. IEEE European Symp. Security and Privacy (EuroS&P)*, Stockholm, Sweden, 2019, pp. 401–415. doi: 10.1109/EuroSP.2019.00037.
8. "Implementing the falsified medicines directive: Safety features," Gov.U.K., 2019. [Online]. Available: <https://www.gov.uk/guidance/implementing-the-falsified-medicines-directive-safety-features>
9. R. Kumar and R. Tripathi, "Traceability of counterfeit medicine supply chain through blockchain," in *Proc. Int. Conf. Communication Systems and Networks (COMSNETS)*, Bengaluru,

- India, 2019, pp. 568–570. doi: 10.1109/COMSNETS.2019.8711418.
10. Using the AWS Blockchain Template for Ethereum, Amazon Web Services. Accessed on: May 10, 2019. [Online]. Available: <https://docs.aws.amazon.com/blockchain-templates/latest/developerguide/blockchain-templates-ethereum.html>
  11. Y. Huang, J. Wu, and C. Long, “Drugledger: A practical blockchain system for drug traceability and regulation,” in *Proc. IEEE Int. Conf. Internet of Things (iThings) and IEEE Green Computing and Communications (Green-Com) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (Smart-Data)*, Halifax, Canada, 2018, pp. 1137–1144. doi: 10.1109/Cybermatics\_2018.2018.00206.
  12. E. Hancock, “The U.K. loses 218 million every year from counterfeit wine and spirits,” *The Drinks Business*, June 2018. [Online]. Available: <https://www.thedrinksbusiness.com/2018/06/the-uk-loses-218-million-every-year-from-counterfeit-wine-and-spirits>
  13. S. Aich, S. Chakraborty, M. Sain, H. Lee, and H. Kim, “A review on benefits of IoT integrated blockchain based supply chain management implementations across different sectors with case study,” in *Proc. Int. Conf. Advanced Communication Technology (ICACT)*, Pyeong Chang, South Korea, 2019, pp. 138–141. doi: 10.23919/ICACT.2019.8701910.
  14. J. Hinckeldeyn and K. Jochen, “Developing a smart storage container for a blockchain-based supply chain application,” in *Proc. Crypto Valley Conf. Blockchain Technology (CVCBT)*, Zug, Switzerland,

## ABOUT THE AUTHORS

**NEETESH SAXENA** is an assistant professor at Cardiff University, United Kingdom, and leads the Cyber and Critical Infrastructure Security Lab. He received a Ph.D. from the Indian Institute of Technology Indore. He was a draft committee member for IEEE standards IEEE SA 1912 and P2795. He is a Senior Member of the IEEE. Contact him at [nsaxena@ieee.org](mailto:nsaxena@ieee.org).

**IEUAN THOMAS** works with Aptitude Software, London. He received a B.Sc. (Hons.) in forensic computing and security from the Department of Computing and Informatics, Bournemouth University. His research interests include blockchain security, Internet of Things security, and digital forensics. Contact him at [i7433979@bournemouth.ac.uk](mailto:i7433979@bournemouth.ac.uk).

**PROSANTA GOPE** is a lecturer in cybersecurity at the University of Sheffield, United Kingdom, and is a member of the Security of Advanced Systems Research Group. Contact him at [p.gope@sheffield.ac.uk](mailto:p.gope@sheffield.ac.uk).

**PETE BURNAP** is a professor of data science and cybersecurity at Cardiff University, United Kingdom. He is director of Cardiff’s NCSC/EPSRC Academic Centre of Excellence in Cyber Security Research. He also leads artificial intelligence for cybersecurity research at Airbus DTO. Contact him at [burnapp@cardiff.ac.uk](mailto:burnapp@cardiff.ac.uk).

**NEERAJ KUMAR** is a professor in the Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology, India. He received a Ph.D. in computer science and engineering from SMVD University, Katra. He is an editorial board member for *International Journal of Communication Systems, Security and Communication*, and *Journal of Networks and Computer Applications*. He is a Member of the IEEE. Contact him at [neeraj.kumar@thapar.edu](mailto:neeraj.kumar@thapar.edu).

2018, pp. 97–100. doi: 10.1109/CVCBT.2018.00017.

15. N. Alzahrani and N. Bulusu, “A new product anti-counterfeiting blockchain using a truly decentralized dynamic consensus protocol,” *Concurrency Computation: Practice Exp.*, vol. 32, no. 12, p. e5232, 2019. doi: 10.1002/cpe.5232.
16. S. Wang et al., “Blockchain-powered parallel healthcare systems based on the ACP approach,” *IEEE Trans. Comput. Social Syst.*, vol. 5, no. 4, pp. 942–950, Dec. 2018. doi: 10.1109/TCSS.2018.2865526.
17. F. Jamil, L. Hang, K. Kim, and D. Kim, “A novel medical blockchain model for drug supply chain integrity management in a smart hospital,” *Electronics*, vol. 8, no. 5, p. 505, 2019. doi: 10.3390/electronics8050505.





# Blockchain for Video Streaming: Opportunities, Challenges, and Open Issues

**Nabajeet Barman, Deepak G C, and Maria G. Martini**, Kingston University

*While industry has adopted blockchain-based video-streaming platforms, other stakeholders could contribute more to grow the technology and support its implementation. This article reviews current blockchain-based video streaming applications and industrial advancements and identifies today's research activities and future opportunities for investigation.*

**W**ithin the past decade, multimedia, in general, and video streaming, in particular, have radically changed the way we consume information and keep ourselves entertained. Video-streaming services are increasingly being used, thanks to recent streaming-video-on-demand (SVoD) services such as Netflix, Hulu, and Amazon Prime Video, which are overtaking traditional TV broadcast services in the United Kingdom.<sup>1</sup> The user

expectation of any content, at any location, and at any time has resulted in the public acceptance and worldwide growth of such SVoD services. Furthermore, live-gaming video-streaming services such as Twitch and YouTube Gaming have also seen tremendous growth, with Twitch alone being the fourth-highest peak Internet traffic generator in the United States, serving more than 15 million active users daily with almost 1 million concurrent users.<sup>2</sup>

It is important to note that the success of any service depends on consumer satisfaction, which can be characterized in terms of the quality of experience (QoE)

perceived by end users. As a result, there recently was a focus shift from quality-of-service (QoS)-based assessment to QoE-based evaluation. Due to the proliferation of such streaming services and the new paradigm of user satisfaction measurement, both academia and industry have put every effort toward the identification, design, and evaluation of the QoE.

During the past few years, distributed ledger technology (DLT), in general, and blockchain, in particular, have gained increasing attention from various organizations worldwide, especially those involved in financial services/banking, automobiles, health care, insurance, public sectors, and education, among others. The digital ledger market for blockchain products and services is anticipated to reach US\$60.7 billion in 2024, up from US\$708 million in 2017.<sup>3</sup> Blockchain itself is still a nascent technology, and much needs to be done to achieve its full potential so that it is essentially accepted by industry, government, and consumers. Blockchain-based video applications have attracted the attention of industry, with streaming applications such as DLive and Livepeer already available in the market. The Gartner Hype Cycle<sup>4</sup> reports that blockchain business for media applications (blockchain in media and entertainment) is still at the very first stage of the “innovation trigger” phase and will take 5–10 years to reach the plateau of productivity as per current projections.

Due to the many prospective advantages provided by blockchain (see the “Blockchain Technology: Features, Advantages, and Shortcomings” section), various research work has been carried out to explore the technology’s application in different domains, from the Internet of Things<sup>5</sup> and addressing security issues in unmanned aerial

vehicle networks<sup>6</sup> to video surveillance systems<sup>7</sup> and secure video storage.<sup>8</sup> However, there has been little focus from academia on identifying the potential research questions and opportunities provided by blockchain for video-streaming applications.

An IBM executive report<sup>9</sup> discusses how blockchain has every potential to change the way media content is currently transmitted, consumed, and paid for. Additional reports from IBM<sup>9</sup> and the Interactive Advertising Bureau<sup>10</sup> discuss how the video-advertising industry, which is plagued by a lack of transparency, high middlemen charges, and closed performance-measurement metrics, can benefit from using blockchain applications (BAs), which provide a highly efficient and transparent advertising platform at a reduced cost. A Joint Photographic Experts Group (JPEG) whitepaper<sup>11</sup> discusses industrial needs, relevant use cases, and functionalities for media blockchain and the potential requirements of ongoing standardization activities within and outside the organization. In Bhowmik and Feng,<sup>12</sup> a distributed and tamper-proof novel watermarking multimedia blockchain framework for content processing and transaction is briefly discussed. Zhaofeng et al.<sup>13</sup> present a blockchain-based digital rights management (DRM) system to help protect content-ownership rights, while Bui et al.<sup>14</sup> detail a blockchain-based platform, Archangel, for protecting video archives. Ghimire et al.<sup>15</sup> introduce a novel video integrity mechanism based on blockchain for protection from tampering.

Detection of deepfake videos and images is a very challenging problem at present, for which blockchain-based solutions are being investigated in works such as those by Hasan and Salah<sup>16</sup> and JPEG.<sup>17</sup> A detailed blockchain-based

video-delivery model using advanced network services chains has been proposed by Herbaut and Negru.<sup>18</sup> The model provides a platform for collaboration among various providers using network service chains. Other works, including Y. Liu et al.<sup>19</sup> and M. Liu et al.,<sup>20</sup> propose decentralized resource allocation for blockchain-based video-streaming applications, using mobile edge computing (MEC) to improve video delivery to the end user. Transcoding still remains a challenging task, especially for live-streaming applications. Toward this direction, work by Liu et al.<sup>21</sup> presents a deep reinforcement-learning-based transcoder-selection framework for blockchain-based video-streaming solutions for more effective transcoding node selection.

Based on the preceding discussion, it is clear that, while there have been numerous works on the potential use of blockchain technology for media and entertainment, an overview that explores current industrial and academic efforts and identifies various potential advantages and opportunities of blockchain, with a focus on video streaming, has been missing, so far. Toward this end, in this article, we present the following:

1. a discussion of various opportunities provided by the use of blockchain technology for video streaming
2. a review of existing blockchain-based video-streaming platforms
3. a summary of ongoing blockchain standardization activities
4. a conceptual framework for blockchain-based video-streaming applications
5. a discussion of technical challenges and open issues.

## BACKGROUND

### Video streaming

Video-streaming platforms, such as Netflix and YouTube, use a centralized client-server architecture, where multimedia content is delivered from a server to the client through the network. Such centralized systems face the problem of “last mile” delivery, where, due to the low link capacity of the access network to the end users, the QoS is significantly affected. The advent of new technologies, such as 5G and MEC, contributes to overcoming such limitations, with some new works exploring the potential of blockchain technology, such as Y. Liu et al.<sup>19</sup> and M. Liu et al.,<sup>20</sup> discussed previously. Still, many challenges remain due to the increasing bandwidth required to meet the demands of the changing entertainment landscape, involving the use of higher frame rates and resolutions, that is, 4K and 8K, as well as newer video formats, for example, high dynamic range, 3D, light field, and point clouds. A possible solution is the decentralized distribution of content through peer-to-peer (P2P) networks, where data are shared by participating nodes (clients). This sharing of bandwidth by network users can help overcome the bottleneck bandwidth issue, resulting in a comparatively cost-effective, efficient, and faster network.

### QoE

During the past two decades, the industry mindset has shifted from a product-centric to a service-centric approach, where users are at the center of the business model. Measuring or estimating the QoE is critical to understanding the success of a service or technology. The adoption of a new

technology by end users is based on the level of consumer satisfaction, which for video-streaming applications can be estimated using various QoE models and metrics. On the other hand, existing services can benefit from using QoE models and metrics, which can be harnessed to estimate and, as a result, improve end-user satisfaction, ultimately leading to increased viewing times and reduced churn.

### Blockchain technology: Features, advantages, and shortcomings

Blockchain is a digital list of records, that is, blocks linked together using cryptographic algorithms. It gained tremendous attention with the introduction and high popularity of Bitcoin, a cryptocurrency based on blockchain technology. Its property of resistance to modification (immutability) makes it a prospective and highly disruptive technology across a wide range of industries. Some of the significant advantages offered by blockchain include faster and more secure transactions, transparency, cost effectiveness (due to the absence of middlemen), traceability, automated actions using smart contracts, and cryptographically sealed protected data, providing security and privacy. All these advantages come with their fair share of shortcomings, the most important of which is scalability. Due to the limited size of the block, the rate of mining is slow. A bigger block size results in slower propagation in the complex chain of blocks. Also, current blockchain technologies can handle a relatively low transaction rate, typically 7/s for Bitcoin and 15/s for Ethereum.

On an elementary level, blockchains can be broadly classified into three categories: public (namely, Bitcoin and Ethereum), which are fully decentralized

and available to all participants; private (Hyperledger and R3 Corda), which are usually centralized and entities require permission from blockchain owners; and hybrid, also called *consortium blockchain* (for example, Dragonchain and Hyperledger Quilt), which combine the advantages of both the public (transparency and security) and private (privacy) blockchains.

Due to the increasing popularity of blockchain, much like platform as a service, software as a service, and infrastructure as a service, the concept of blockchain as a service (BaaS) is becoming popular. Here, BaaS enables both companies and consumers to develop, test, and deploy secure blockchain apps with the required functionality, such as bandwidth management, resource allocation, and hosting requisites. Currently, BaaS is provided by major companies, such as IBM, Microsoft, and Oracle, and it is increasingly being accepted by various industries because it enables them to focus on their product/service offerings without having to worry about underlying blockchain network protocols and reference models. For a detailed overview of blockchain's fundamental concepts, see, for example, Puthal et al.,<sup>27</sup> for its possible impact on society, see, in particular, Aste et al.<sup>28</sup> and the references therein.

## BLOCKCHAIN FOR VIDEO STREAMING

During the past few years, the world has seen a proliferation of smartphone users, which has dramatically changed the way we produce and consume digital content. People have taken on the role of content producer, from recording/live-streaming events such as concerts and film festivals to video blogs, for example, travel, tutorials, and product reviews, and





**FIGURE 1.** The opportunities offered by blockchain technologies for video-streaming applications.


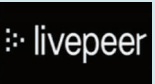




live-gaming video streaming. Limited by the shortcomings of traditional platforms, which fail to adequately reward content creators and viewers, we have seen, in recent years, the rise of platforms such as Medium, Bonzo Me, and Tsu, which reward content producers and, in some cases, viewers, leading to a win-win environment for all parties. One of the much-anticipated use cases for blockchain-based video-streaming platforms is, indeed, in the domain of live streaming and gaming/e-sports. Some of the major prospective advantages offered by

blockchain-based video-streaming applications include empowering creators and artists, blockchain-based content-delivery networks (CDNs), privacy and piracy protection, and energy saving, as presented in detail in Figure 1.

To shed light on the current market scenario, the state of technology adoption, and various challenges currently faced by the industry, we present in Table 1 a brief comparative review of six of the biggest companies providing blockchain-based video-streaming solutions. We can

observe that blockchain-based video streaming has already been introduced in the marketplace. In addition to these six companies, Play2Live, the world's first full-scale blockchain-based decentralized streaming service with its own blockchain platform, called *Level Up Chain*, needs to be mentioned. With a major focus on the e-sports domain, Play2Live provided additional services, such as betting and gambling. Due to regulatory challenges, the company ceased operations at the end of 2018, and its fate at the time of writing this article

**TABLE 1.** Blockchain-based video-streaming companies.

Company	Blockchain platform	Digital currency	Primary application domain	Key characteristics
	Lino blockchain (third party)	Lino points	Live streaming	<ul style="list-style-type: none"> <li>• No commission</li> <li>• Engagement bonus for viewers (for watching, chatting, and gifting)</li> <li>• Content streamers/producers get 90.1% of the money and an additional content bonus.</li> <li>• The remaining 9.9% of the money rewards people for their contributions to the network on a daily basis.</li> </ul>
	Ethereum blockchain (third party)	Ether and Livepeer tokens (LPTs)	Video streaming and transcoding	<ul style="list-style-type: none"> <li>• Highly scalable, low-cost, decentralized CDN, point-to-point-based streaming application</li> <li>• Two major stakeholders are transcoders (different bitrates and/or packaging formats) and delegators (token holders who staked their tokens to the transcoders).</li> <li>• Delegators can earn rewards/tokens for their work as a share of the money paid by the broadcaster and from newly mined LPTs.</li> </ul>
	Theta blockchain (proprietary)	Theta token	Generic (e-sports, music, TV, and so on)	<ul style="list-style-type: none"> <li>• Two major components: open-source Theta network and platform and Theta tokens</li> <li>• Supports many different kinds of content streaming (for example, e-sports, music, and TV)</li> </ul>
	VideoCoin Network	VideoCoin	Live streaming	<ul style="list-style-type: none"> <li>• Fuji, released in August 2019, is the first product based on the VideoCoin Network.</li> <li>• Enables live streaming and future releases to include support for on-demand video encoding, transcoding, and distribution</li> </ul>
	BitTorrent plus Ethereum	Flixx	Video distribution	<ul style="list-style-type: none"> <li>• Combines P2P payments and P2P content distribution to make a social experience</li> <li>• Content creators can distribute their products straight to end users and get paid directly.</li> </ul>
	LBRY network	LBRY credits (LBCs)	Content sharing	<ul style="list-style-type: none"> <li>• Available on all platforms (apps and web), it supports sharing videos, music, games, e-books, and so on.</li> <li>• Open source platform that uses LBCs for publishing, purchasing, and supporting content on the network.</li> </ul>

remains unknown. This indicates the need for international standards guided by technology and regulations so that blockchain technology and its applications can grow further and gain increasing acceptance worldwide. We present in Figure 2 a summary of current standardization activities.

### CONCEPTUAL FRAMEWORK FOR BLOCKCHAIN-BASED VIDEO-STREAMING APPLICATIONS

Based on the previous discussion of the need for unifying various ongoing works to achieve better, more efficient, and more interoperable systems, it is essential to design a general

architecture describing multiple network elements and modules and how they interact with new and existing systems. Toward this end, we present in Figure 3 a conceptual framework for blockchain-based video-streaming applications, considering the whole multimedia delivery chain, from content production and transmission to

## BLOCKCHAIN: FROM TECHNOLOGY TO MARKETPLACES



FIGURE 2. The standardization activities.

consumption. We follow an on-the-top modular approach based on the concept of BaaS, which offers the advantage that, depending on the streaming requirements, the BA can be run with newly designed networks, by

integration into existing applications, or both.

Due to the low-cost, open, and high-scalability advantages offered by distributed, decentralized applications, there is already an effort toward the design

of a decentralized web. For example, in the case of CDNs, some existing centralized solutions, such as Cloudflare, have started offering a decentralized content gateway via the InterPlanetary File System (IPFS), which can be used

by blockchain-based applications for decentralized file storage and access. Also, tools such as Hyperledger Quilt offer interoperability between ledger systems for payments and transfer value across distributed and undistributed ledgers.

However, there is still a lack of technological capabilities for supporting video streaming in a fully decentralized manner. Hence, the proposed conceptual framework is designed to offer a modular, highly scalable design, where the application developer has the flexibility to decide on the exact type of solution (centralized, decentralized, or distributed) to be used at each stage of multimedia delivery, depending on the requirements. We discuss next the three major roles one can select at the different stages of the multimedia delivery chain, as shown in Figure 3.

1. The broadcaster is an over-the-top service provider or a user who wants to deliver a video through the network to viewers. To encode and/or transcode a video along with encryption and packaging, the broadcaster can use existing centralized solutions (for example, cloud-encoding solutions, such as Amazon Web Services) or decentralized transcoding solutions, where various entities act as transcoding nodes to perform these tasks in return for payment through digital currency (namely, Bitcoin, Ether, and so on).
2. The CDN can be centralized (for example, Akamai), decentralized (such as VideoCoin and Theta Token), or a combination of both, as required.
3. The broadcaster can choose to use a distributed/decentralized network of relay nodes

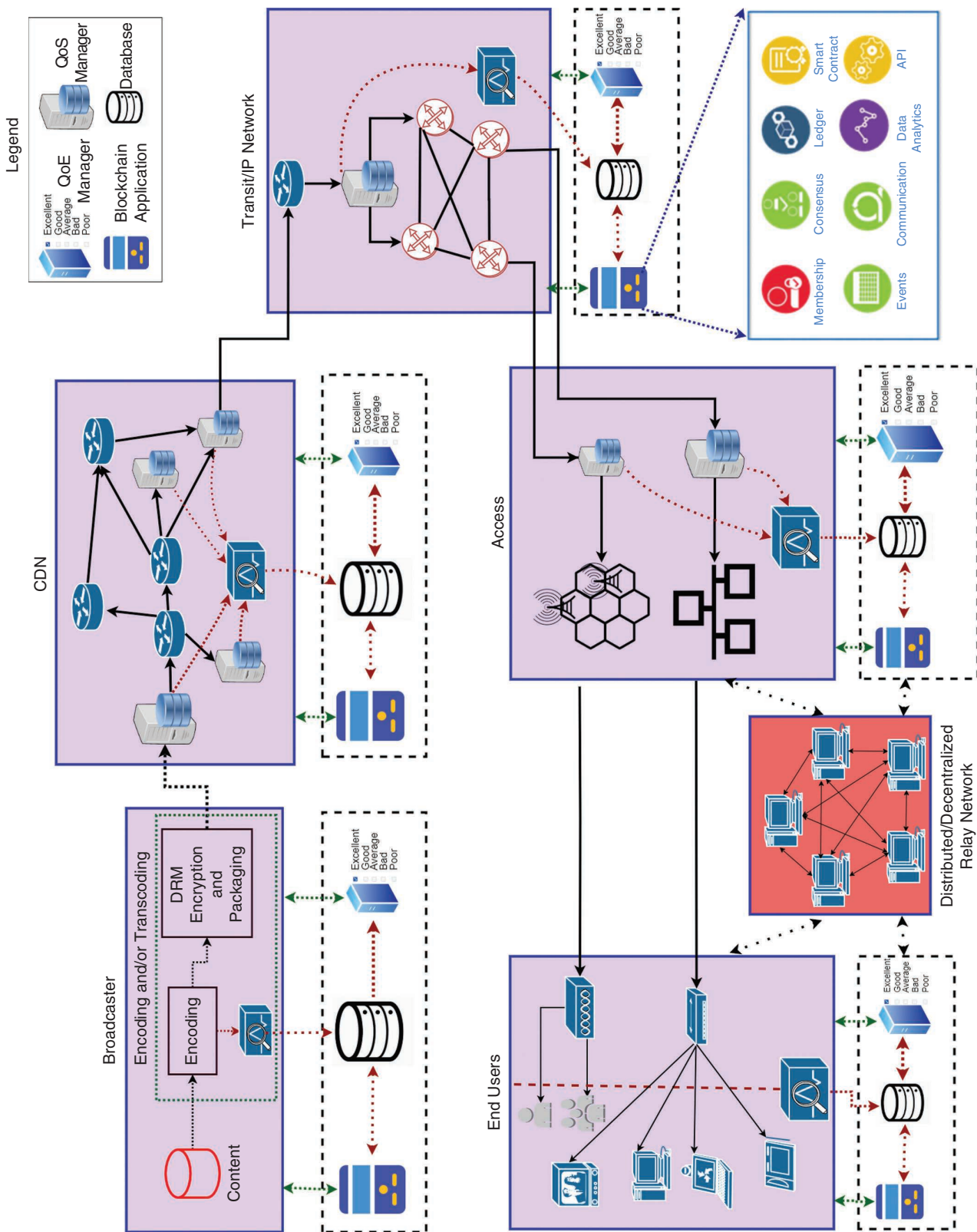
to assist with distribution of a video. Such relay nodes can help reduce network latency as well as provide high scalability. Music-streaming applications, such as Spotify, use both the P2P network and centralized client-server architecture to deliver the fastest possible service to end users. Like transcoding nodes, relay nodes are paid for the tasks they perform, using digital currencies. Finally, videos are delivered to end users, who pay the broadcaster for the service.

The proposed architecture contains four additional modules: the BA, QoE server, QoS server, and database. The BA is the core module implementing the blockchain framework. Depending on the application requirements, it can be a permissioned blockchain (for example, Hyperledger), permissionless (Ethereum), or hybrid. Depending on the stage of the multimedia delivery and type of underlying network (centralized/decentralized), the BA will consist of one or more of the following different functional groups (as shown in Figure 3):

1. membership, for example, participating nodes and their identifications, roles, rights, and permissions (such as Hyperledger Composer on top of Hyperledger Fabric)
2. a ledger that contains metadata, such as timestamps; video content-related information (such as recording and publishing rights and ownership information); and smart contract payment information (in particular, chaincode in Hyperledger and Ethereum)

3. consensus algorithms, for example, proof of work and proof of stake
4. smart contracts, which can be automatically executed at the end of an event/completion of a transaction (for example, in the proposed framework, automated payments such as from the broadcaster to the transcoding and/or relay nodes and from the end user to the broadcaster for using the service, can be automatically triggered using prenegotiated terms programmed into a smart contract)
5. application programming interfaces (APIs) to enable the BA to connect to and interact with different network elements, such as representational state transfer APIs
6. data analytics, which can provide overall insight into the functioning of various components, such as QoE values, the number of successful transactions, the number of nodes, and so forth, supplemented by various visualization and analysis tools
7. a communication submodule to implement the connection protocols that are used and the message type, for example, based on the Transmission Control Protocol/User Datagram Protocol, as in standard web applications, or an IPFS-based decentralized protocol for distributed storage communication among various nodes
8. the events submodule to detect different interactions, such as video transmission, displays, user interactions,





**FIGURE 3.** The proposed conceptual framework for blockchain-based video streaming (the icons of the BA components are borrowed from the IBM blockchain reference architecture template).

and other events occurring during streaming; event listeners and message buses can be used to accomplish the same.

QoS and QoE servers play an important role in the measurement and monitoring of various network and user experience parameters to make sure that smart contract terms based on various service-level agreements (SLAs) and/or experience-level agreements (ELAs) are adhered to by participating entities. The database stores all information required by the network. It can be a centralized, horizontally scalable (MongoDB) database; a vertically scalable (Structured Query Language) one; or a decentralized data-storage application, such as Filecoin or Swarm.

## TECHNICAL CHALLENGES AND OPEN ISSUES

1. *Design of a scalable blockchain-based video-streaming platform:* Video streaming, in general, consists of a large number of sessions, which result in multiple transactions during a single stream. This may lead to a remarkably high number of transactions, which can produce long delays and/or high computational complexity. Therefore, a development platform is critically required to implement innovative strategies in a custom blockchain platform/model for video streaming to overcome the existing challenges of longer transaction times and limited computing power. Current efforts in this direction, for instance, from Play2Live, which used the Graphene platform

(third-generation software with cryptographic abilities capable of performing 50,000 transactions per second), can be further explored.

2. *Payments and revenue-sharing model:* An exciting and much-needed research challenge is the design of revenue models that are fair and maximize the benefits for all involved stakeholders, in particular, content producers, publishers, aggregators, technology providers, advertisers, and legal-service providers. At present, artist contracts and payments are severely limited due to the presence of such intermediaries. Blockchain with a smart contract can help design an automated payment system, which ultimately enables artists and creators to get their rightful shares and retain ownership of their content. Along with the revenue model, the design of proper SLAs and ELAs is required to assign tasks and responsibilities to each stakeholder.
3. *Protection of content owners' rights:* A thorough investigation is required to develop specific provisions and features to help track ownership and prevent illegal copying and distribution of content. Application-specific smart contracts incorporating licensing and ownership-related features will need to be developed. These could help address DRM issues, which remain an enormous challenge for the media industry.

4. *Lack of standards and regulation:* Currently, one of the most significant disadvantages of blockchain technology is uncertainty about the legal status of its applications, such as cryptocurrency, betting, and gambling. For the technology to mature and be accepted by different public and private entities, we need an international blockchain-based video-streaming standardization group, including both technical and legal experts representing industry, academia, and regulators. There is a need to develop video-streaming-related smart contract standards similar to the Ethereum Request for Comments-20 Ethereum Smart Contract. Other standards to maintain interoperability between legacy and newly developed blockchain-based video-streaming platforms, such as the Interledger Protocol, also need to be developed.
5. *Integration and interoperability:* To make sure of interoperability with legacy systems, more research toward development of standard protocols and terminology is required. Also, new network technologies, such as fog computing, MEC, software-defined networking, and network functions virtualization, are being introduced. Therefore, the designed solution should also be interoperable with newly developed technologies and protocols. Also, with the rise of multiple decentralized streaming companies, there is a need for

## ABOUT THE AUTHORS

**NABAJEET BARMAN** is a research associate at Kingston University, London. His research interests include multimedia communications, machine learning, and, more recently, distributed ledger technologies. Barman received a Ph.D. from Kingston University. He is a Video Quality Expert Group board member, and he is involved in ITU-T standardization activities and an MSCA fellow. He is a Member of the IEEE. Contact him at [h2o.is.water@gmail.com](mailto:h2o.is.water@gmail.com).

**DEEPAK G C** is a lecturer at the School of Computer Science and Mathematics, Kingston University, London. His research interests include radio-access technologies, cognitive radio, 5G, physical layer security, the Internet of Things, cybersecurity, and public safety communications. He received a Ph.D. from Lancaster University. He is a Member of the IEEE. Contact him at [d.gc@kingston.ac.uk](mailto:d.gc@kingston.ac.uk).

**MARIA G. MARTINI** is a professor in the Faculty of Science, Engineering, and Computing, Kingston University, London. Martini received a Ph.D. in electronics and computer science from the University of Bologna, Italy. She has been an associate editor for *IEEE Signal Processing Magazine* and *IEEE Transactions on Multimedia* and a vice chair of the IEEE MMTC. She currently chairs IEEE Standard WG P3333.1.4. She is a fellow of the Higher Education Academy and a Senior Member of the IEEE. Contact her at [m.martini@kingston.ac.uk](mailto:m.martini@kingston.ac.uk).

standardization across various ledgers and tokens to ensure interoperability with different currencies/tokens on the blockchain.

6. *Big data analytics for machine learning (ML)/artificial intelligence (AI)*: Data and information stored by the BA can be used for data analytics. Blockchain technology with inherent data security features can provide additional functionality for automated tasks, such as generating reports based on different ML and AI algorithms. Computation can be performed more efficiently due to the availability of the

joint power of the participating nodes. More efforts in this direction can help us understand and improve network efficiency and user satisfaction.

7. *QoE measurement, modeling, and control*: QoE remains an integral part of today's video-streaming applications. Since a blockchain-based video-streaming platform is different from a traditional centralized video-streaming platform as presently used, the applicability and efficiency of currently developed models for these newer applications remain open

questions. Therefore, it is necessary to identify various influencing factors, such as QoE monitoring and control models, among others, that are designed for such applications. Furthermore, identifying such factors is required for successfully enforcing SLAs and ELAs between various stakeholders.

In this article, we discussed blockchain technology and its potential role in and advantages for video-streaming applications. We also investigated various applications that are already implemented in blockchain video streaming and pointed out that there exists a significant gap between industry and academia. Standardization will be an important enabler of the adoption of blockchain for media streaming. To address this aspect, we presented a discussion of relevant activities in various standardization groups. We also discussed a conceptual framework, highlighting the possible network architecture and interaction between various elements. Based on our work, we identified key technical challenges and open research questions, which will help bring the technology to the marketplace successfully. ■

## ACKNOWLEDGMENTS

We acknowledge the support of the European Commission project Horizon 2020-643072, Innovative QoE Management in Emerging Multimedia Services.

## REFERENCES

1. "TV streaming services overtake pay TV for first time," Ofcom, London. July 17, 2018. [Online]. Available: <https://>

- www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2018/streaming-overtakes-pay-tv
2. N. Barman, "An objective and subjective quality assessment for passive gaming video streaming," Ph.D. dissertation, Kingston Univ., London, 2019.
3. "Blockchain market shares, market strategies, and market forecasts, 2018 to 2024," IBM, Armonk, NY. Jan. 2018. [Online]. Available: <https://www.ibm.com/downloads/cas/PPRR983X>
4. H. P. Levy, "The reality of blockchain," Gartner, Stamford, CT. Oct. 22, 2018. [Online]. Available: <https://www.gartner.com/smarterwithgartner/the-reality-of-blockchain/>
5. M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of blockchains in the Internet of Things: A comprehensive survey," *IEEE Commun. Surveys Tut.*, vol. 21, no. 2, pp. 1676–1717, 2nd quarter 2019. doi: 10.1109/COMST.2018.2886932.
6. J. Qiu, D. Grace, G. Ding, J. Yao, and Q. Wu, "Blockchain-based secure spectrum trading for unmanned-aerial-vehicle-assisted cellular networks: An operator's perspective," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 451–466, Jan. 2020. doi: 10.1109/JIOT.2019.2944213.
7. Y. Jeong, D. Hwang, and K. Kim, "Blockchain-based management of video surveillance systems," in *Proc. Int. Conf. Information Networking (ICOIN)*, Jan. 2019, pp. 465–468. doi: 10.1109/ICOIN.2019.8718126.
8. S. S. Arslan and T. Goker, "Compress-store on blockchain: A decentralized data processing and immutable storage for multimedia streaming. 2019. [Online]. Available: arXiv:1905.10458
9. "Enforcing accountability in media: How blockchain technology can work for media and entertainment," IBM, Armonk, NY. Sept. 2018. [Online]. Available: <https://www.ibm.com/downloads/cas/6146Z4JE>
10. "Blockchain for video advertising: A market snapshot of publisher and buyer use cases," IAB, New York. Feb. 2018. [Online]. Available: [https://www.iab.com/wp-content/uploads/2018/02/Blockchain\\_for\\_Video\\_Advertising\\_Publisher-Buyer\\_Use\\_Cases\\_2018-02.pdf](https://www.iab.com/wp-content/uploads/2018/02/Blockchain_for_Video_Advertising_Publisher-Buyer_Use_Cases_2018-02.pdf)
11. "JPEG white paper: Towards a standardized framework for media blockchain and distributed ledger technologies," White Paper, Aug. 2019. [Online]. Available: <https://jpeg.org/static/whitepapers/jpeg-media-blockchain-whitepaper.pdf>
12. D. Bhowmik and T. Feng, "The multimedia blockchain: A distributed and tamper-proof media transaction framework," in *Proc. 22nd Int. Conf. Digital Signal Processing (DSP)*, Aug. 2017, pp. 1–5. doi: 10.1109/ICDSP.2017.8096051.
13. M. Zhaofeng, H. Weihua, and G. Hongmin, "A new blockchain-based trusted DRM scheme for built-in content protection," *EURASIP J. Image Video Process.*, vol. 2018, pp. 11,169–11,185, Sept. 2018. Art. no. 91.
14. T. Bui et al., "Tamper-proofing video with hierarchical attention autoencoder hashing on blockchain," *IEEE Trans. Multimedia*, to be published. doi: 10.1109/TMM.2020.2967640.
15. S. Ghimire, J. Y. Choi, and B. Lee, "Using blockchain for improved video integrity verification," *IEEE Trans. Multimedia*, vol. 22, no. 1, pp. 108–121, Jan. 2020. doi: 10.1109/TMM.2019.2925961.
16. H. R. Hasan and K. Salah, "Combating deepfake videos using blockchain and smart contracts," *IEEE Access*, vol. 7, pp. 41,596–41,606, 2019. doi: 10.1109/ACCESS.2019.2905689.
17. "JPEG Committee releases a call for evidence for image compression based on AI," JPEG, Geneva, Switzerland, Feb. 17, 2020. [Online]. Available: [https://jpeg.org/items/20200217\\_press.html](https://jpeg.org/items/20200217_press.html)
18. N. Herbaut and N. Negru, "A model for collaborative blockchain-based video delivery relying on advanced network services chains," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 70–76, Sept. 2017. doi: 10.1109/MCOM.2017.1700117.
19. Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Decentralized resource allocation for video transcoding and delivery in blockchain-based system with mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11169–11185, Nov. 2019. doi: 10.1109/TVT.2019.2937351.
20. M. Liu, F. R. Yu, Y. Teng, V. C. M. Leung, and M. Song, "Distributed resource allocation in blockchain-based video streaming systems with mobile edge computing," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 695–708, Jan. 2019. doi: 10.1109/TWC.2018.2885266.
21. Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung, "Decentralized resource allocation for video transcoding and delivery in blockchain-based system with mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11,169–11, 185, Nov. 2019. doi: 10.1109/TVT.2019.2937351.
22. "Twitch affiliate agreement," Twitch, San Francisco, CA, Aug. 10, 2018. [Online]. Available: <https://>



- www.twitch.tv/p/en-gb/legal/affiliate-agreement/
23. "Payments and earnings," Uber, San Francisco, CA, 2019. [Online]. Available: <https://www.uber.com/en-gh/drive/resources/payments/>
  24. C. Cadwalladr and E. Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach," *The Guardian*, Mar. 17, 2018. [Online]. Available: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
  25. Based on a true story. Flixxo, Buenos Aires, Argentina, Sept. 6, 2019. [Online]. Available: <https://medium.com/@flixxo/based-on-a-true-story-12e8e8b51bc0>
  26. J. Trunz, "VideoCoin analysis," Oct. 17, 2018. [Online]. Available: <https://medium.com/dcryptinc/video-coin-analysis-a3d433b9a126>
  27. D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and G. Das, "Everything you wanted to know about the blockchain: Its promise, components, processes, and problems," *IEEE Consum. Electron. Mag.*, vol. 7, no. 4, pp. 6–14, July 2018. doi: 10.1109/MCE.2018.2816299.
  28. T. Aste, P. Tasca, and T. D. Matteo, "Blockchain technologies: The foreseeable impact on society and industry," *Computer*, vol. 50, no. 9, pp. 18–28, Sept. 2017. doi: 10.1109/MC.2017.3571064.



IEEE COMPUTER SOCIETY

**DIGITAL LIBRARY**

Access all your IEEE Computer Society subscriptions at

**computer.org**  
**/mysubscriptions**

**SUBMIT  
TODAY**

IEEE TRANSACTIONS ON  
**BIG DATA**

► **SUBSCRIBE AND SUBMIT**

For more information on paper submission, featured articles, calls for papers, and subscription links visit: [www.computer.org/tbd](http://www.computer.org/tbd)

*TBD* is financially cosponsored by IEEE Computer Society, IEEE Communications Society, IEEE Computational Intelligence Society, IEEE Sensors Council, IEEE Consumer Electronics Society, IEEE Signal Processing Society, IEEE Systems, Man & Cybernetics Society, IEEE Systems Council, and IEEE Vehicular Technology Society

*TBD* is technically cosponsored by IEEE Control Systems Society, IEEE Photonics Society, IEEE Engineering in Medicine & Biology Society, IEEE Power & Energy Society, and IEEE Biometrics Council



**IEEE**



IEEE  
COMPUTER  
SOCIETY



# Blockchain for E-Health-Care Systems: Easier Said Than Done

**Sujit Biswas, Kashif Sharif, and Fan Li**, Beijing Institute of Technology

**Saraju P. Mohanty**, University of North Texas

*Blockchain as a distributed ledger technology can be very effective in providing access control and big data management in health-care systems. Because implementing or migrating to a pure blockchain solution is an extremely challenging task, several design and implementation dynamics should be considered.*

**B**lockchain has become a popular buzzword in recent years, giving the impression that it is a “silver bullet” to several (if not all) security problems. There is no denying that it does make systems more transparent, traceable, and secure; however, it is in no way a one-solution-fits-all technology. One can easily find several research works focused on using blockchain in different applications, ranging from industrial automation to vehicular networks and from the Internet of Things (IoT) to financial markets. From a practical

perspective, these are easier said than done. In this article, we focus on the health-care industry and explore how it can benefit from distributed ledger technology (DLT) in general, and blockchain specifically, for process automation, digital/electronic medical record (EMR) management (including big data), access control, and smart contracts (SCs). Several works in literature have focused on these topics;<sup>1,2</sup> however, most of them solve very specific challenges while ignoring the related bigger picture.

## **DON'T THINK CRYPTOCURRENCY**

Blockchain has gained significant attention in the past few years primarily due to the skyrocketing prices of

Bitcoin. Since then, dozens of cryptocurrencies have sprung up around the globe. Perhaps the biggest misconception about blockchain is that it is for cryptocurrencies only. Blockchain is primarily a DLT but has mostly been specialized for financial transactions. However, generic DLT mainly focuses on providing a set of protocols and processes for the distribution of records among multiple nodes in a collaborating system.<sup>3</sup> The system may belong to a single enterprise, or multiple enterprises may connect to a single, yet shared, DL. Thus, blockchain inherits the benefits of DLTs and adds a few more to the list, such as 1) security through SCs, which are predefined agreements among parties to conduct business, 2) transparency and accountability through immutable records stored at distributed locations, and 3) efficiency and cost reduction due to the automation of processes.

Blockchains for cryptocurrencies revolve around the concept of tokens that are exchanged among participating users. However, the benefits offered are not limited to tokens. Consider the token as a data element that is generated and traded while leaving an audit trail behind; then, any digital asset (or piece of information) transferred among participants while requiring an audit trail can potentially benefit from blockchain. Besides, access control for such digital assets can be efficiently implemented through SCs while the data itself can be stored in the DL system, increasing its reliability and authenticity. Based on these arguments, the use of blockchain beyond cryptocurrencies is not only feasible but also very practical.

Business blockchain (BBC), a variant of traditional blockchain, aims at using the protocols of blockchain

within a business process, such as the collection of authenticated and verified data from assembly line sensors, the casting and auditing of votes in an e-government solution, or asset tracking.<sup>4,5</sup> Another method used for the classification of blockchain is based on the openness of the system, that is, public, consortium/federated, or private blockchains. Public blockchains are open to all, while consortium blockchains are limited to a group of organizations and private to a specific organization. Public blockchains have publicly open access, and anyone can become a miner, peer, or trader. Contrarily, consortium/private blockchains usually have permissioned access, where users are first registered and authenticated. BBC is usually consortium/private with permissioned access, where a peer is responsible for verification consensus formation.<sup>6</sup>

Hyperledger<sup>7</sup> is a Linux foundation solution that can be used as a base platform for implementing BBCs; hence, most of the debate in this article involves its use and the flexibility it offers. It implements five frameworks intended for different types of environments and consensus mechanisms. Hyperledger Fabric is a major implementation that enables flexible consensus algorithm implementation, SC integration, and IoT support. It is important to note that Hyperledger Fabric is only a platform and does not provide a complete business solution for blockchain in any specific scenario.

### HEALTH-CARE BLOCKCHAINS

The digitization and integration of the IoT in e-health-care systems (eHSs) has made it one of the fastest-growing domains, thus evolving to smart health care.<sup>8</sup> Statistics show that

global health-care spending will continue to rise in 2020 and beyond, with a significant emphasis on digital transformation.<sup>9,10</sup> Medical service providers will increase the use of innovative solutions such as cloud computing, 5G, big data analytics, blockchain, artificial intelligence, and so forth to reduce costs and improve the quality of care.

Integrating blockchain with eHSs can have several benefits, including the security of EMRs, access control for different types of users, the automated execution of services, remote data collection and logging, the unification or standardization of information, redundancy and fault tolerance, the enforcement of health-care regulations, logistics, and so on.<sup>11-13</sup> However, realizing such a blockchain is extremely challenging. To begin with, a modern eHS is a combination of many different technologies at the device level as well as at the operational and management system levels. Hence, the blockchain solution should not only cater to the needs of small-scale sensor devices but also accommodate the devices that generate heavy images [computed tomography (CT) scans]. At the same time, these data have to be shared across departments and with third-party service providers, such as insurance companies. To further complicate things, interoperability among different service providers may not be possible at all due to completely different automation solutions.

To be more specific, some of the major challenges can be listed as follows:

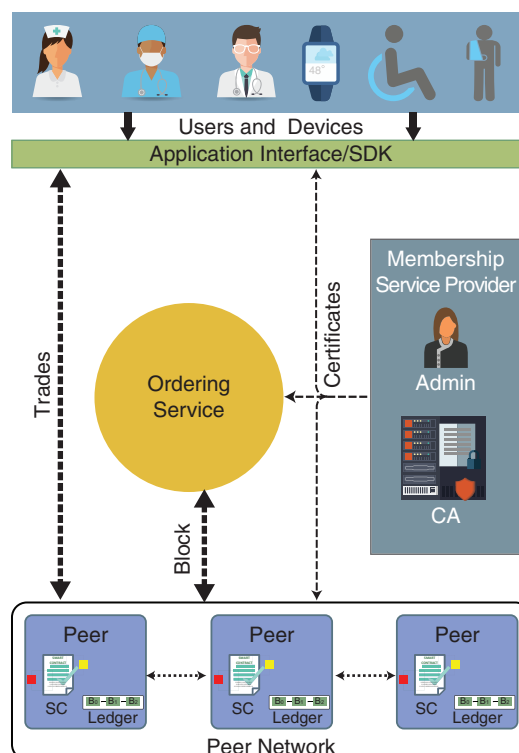
1. Existing centralized eHSs store data in relational databases (DBs), whereas blockchain uses a file DB, and DB schema may not have a one-to-one mapping.

2. Due to restrictions on transaction size in a block, it is impossible to store complete medical imagery as part of the chain.
3. Due to real-time transactions at a mass scale, it is challenging to migrate all the medical histories of patients to a blockchain ledger.
4. In an eHS, it is possible that some medical documents are paper based; hence, the only way to digitize them is to store them as images, which is a nonreal-time process.
5. Access to patient data has to be tightly regulated for different types of internal and external users.<sup>14</sup>
6. An eHS may allow for the integration of third-party IoT devices (smartwatches and health sensors) to be a part of the system, which makes verification and validation difficult.
7. Interaction with other non-blockchain subsystems of the e-health-care ecosystem, including regulatory bodies, may prove challenging.

This is a nonexhaustive list of major challenges that arise when designing a complex blockchain solution for eHSs. In the following sections, we elaborate on every aspect of designing such a system and debate on the technical aspects of different solutions. It is important to note that the objective of this article is not to propose a complete solution but to enable the reader to understand what the challenges are and what the benefits of different possible solutions

can be, although we are inclined toward specific design choices.

A generic BBC process is depicted in Figure 1. Users generate trades (transactions) containing digital assets that need to be shared with other users or devices. A membership service provider comprises an administrator and certificate authority (CA) responsible for providing keys, signatures, certificates, and configuration information. Peers are specialized nodes with resources to execute consensus algorithms and maintain the DL. An ordering service is responsible for grouping all the endorsed/approved trades into a newly generated block. An SC or chaincode is deployed on the peer nodes for verifying transaction agreements among different users.



**FIGURE 1.** A BBC operational framework. SDK: software development kit.

## ELECTRONIC HEALTH RECORD PRIVACY AND REGULATIONS

The primary reason to integrate blockchain into any system is the enhancement of security. It is important to understand that blockchain only adds validation and immutability to the asset-exchange process and stored data, respectively. However, these additions have a significant and profound impact on the overall security architecture. The validation is done through SCs and a consensus protocol that ensure no illegal exchange happens, while immutability is achieved by hash connectivity in the chain, ensuring that nothing is changed afterward. Blockchain does not introduce any new encryption algorithm, signature mechanism, hash function, and so forth; hence, the efficient use of existing or development of new algorithms in this regard is extremely important.<sup>15</sup> In an e-health-care use case, several security and privacy primitives need to be reconsidered. For example, some blockchain systems (with public miners) allow the miners to read transaction payload for validation and SC execution. In e-health care, this payload can be an electronic health record (EHR), which must not be shared (even in encrypted format). In a private blockchain, the compromised (or colluding) miner/peer cannot be ruled out; accordingly, the privacy of EHRs may be compromised. Similarly, if the digital signatures used for validation can be linked to specific patients or their physicians, this may also contribute to a breach of privacy and health-care regulations.



Based on this, two things must be considered before designing a blockchain-based eHS.

1. *Understanding regulations:* The Health Insurance Portability and Accountability Act (HIPAA)<sup>16</sup> and the General Data Protection Regulation (GDPR)<sup>17</sup> must be followed, and misconceptions about both must be removed. For example, many researchers equate GDPR with the right to forget; however, the regulation clearly states that for medical practitioners this is not an absolute right. Hence, data privacy as specified by HIPAA/GDPR must be enshrined in the system, for both internal and external elements.
2. *Identifying the blockchain use case:* In light of privacy and regulations, it is imperative that the use of blockchain within the health-care system must be identified. For example, consent management is a cornerstone of health-care regulations; blockchain can be efficiently used for it. Similarly, access control to EHR, drug control, prescription administration, patient monitoring, insurance, accounting, and so on, where immutability and accountability are necessary, can significantly benefit from blockchain.

Blockchain cannot be considered a blanket solution for all health-care subsystems. The enforcement of regulations such as GDPR and HIPAA will be best done through SCs. As a result, mechanisms are needed to guarantee that SCs are written in such a way as to ensure that privacy

regulations are met with respect to health care.

### MIGRATION ISSUES

Designing and implementing any new system for a large- or medium-scale organization always require crossover time with the old system. Slowly, the old system is phased out while data and operations are migrated to the new system. Most of the research in blockchain focuses on algorithmic technicalities and disregards the fact that the initialization time for a blockchain system, especially for health-care organizations, may render the new solution infeasible. In the following sections, we approach this challenge from two aspects.

#### Infrastructure and architecture changes

Traditional eHSs are usually centralized, as shown in Figure 2. The central application, its associated DB, and perhaps the CA are all hosted on a single server. The server may be in the cloud, but from an implementation perspective, it is still a centralized system. It is also possible that a large-scale eHS provider has diversified locations and thus has many centralized systems that collaborate at different levels. This creates an entirely different architecture, as the DBs may be distributed while the web-based application may be centralized.

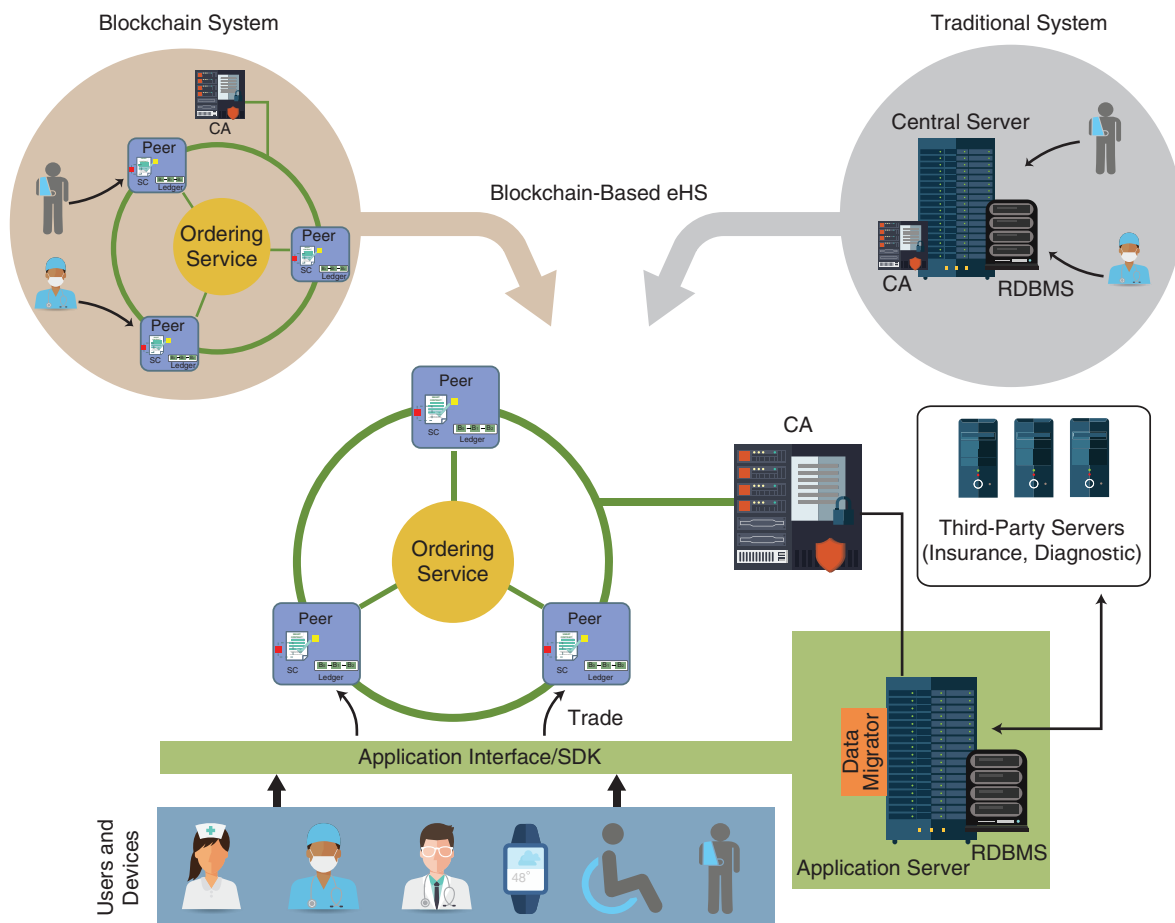
Compared to this, the blockchain system is entirely decentralized. Furthermore, this decentralization is not similar to those of decentralized DBs or distributed systems. As shown in Figure 2, the collection of peer nodes forms a special peer network, which performs consensus formation, while a specialized ordering service (a collection of orderer nodes) is responsible for block formation and its dissemination

back to the peer network. Shifting from a centralized to a distributed blockchain requires significant changes in the infrastructure. This challenge has to be taken into account when designing solutions.

It is important to note that neither users nor devices can initiate trades without an application interface. Many systems employ thin clients on the user side, which means that there has to be an application server as part of the blockchain network. Most research works trivialize this aspect, and users are assumed to be sending the trades directly to the peer. However, in reality, it may be the application server that does it. For thick clients, this assumption may be completely safe, but user devices would still interact with a system entity that manages access control. Figure 2 shows what a combined system would look like. It can be intuitively observed that the application server can create a single point of failure; as a result, it is important to remember that just using blockchain does not make a system tamper-proof. This leads to several new challenges for the securing and interfacing of blockchain with other systems, where secure and standardized application programming interfaces for system interaction should be developed.

#### Data synchronization and migration

One of the least researched areas of blockchain implementation is the migration of existing records and DBs to the new system. Perhaps the simplest reason for this is that the migration of records in their current form is not possible. First, the ledger is unable to accept the previous record with old timestamps. Every new transaction must have a current timestamp. Second, the blockchain ledger is immutable, which means that any



**FIGURE 2.** A blockchain-based integration framework. RDBMS: relational DB management system.

timestamp change after block creation is impossible. Hence, any adjustment or updating has to be done before the migration starts. This is a nontrivial task and may change from one eHS to the other. Third, the traditional centralized system may have thousands of records for hundreds of patients. To bootstrap the blockchain system with all that data at initialization time can be a very long process. All the while, the same system might be in use and creating (or possibly changing) the existing data. This creates a circular migration issue, which must be addressed at

design time. Moreover, efficient migration algorithms and synchronization techniques are needed for this purpose.

One possible solution is to migrate the data not at the initialization phase but only on an as-needed basis. As shown in Figure 2, a data migrator module can be used for formatting the relational DB records into ledger-acceptable trades, only when needed. For example, a patient who has been visiting an eHS has multiple records in the traditional system. After the blockchain migration, when the same patient visits the facility, only then are the necessary records

synchronized. All of the new records are made in the blockchain system, while the relational DB is only used as an old repository. This will ensure that the circular record updating is avoided and initial bootstrap time is negligible. Efficient designs for such data-migration interfaces and algorithms will be the key to successful migration.

## UNIFIED OR MULTIPLE BLOCKCHAINS

Blockchain solutions must be application specific. In an e-health-care scenario, there can be multiple service

providers each with their own independent systems. Cooperation among these systems can be enabled if there is an operational-level agreement. However, transferring the EMRs of a single patient to one another or unifying them in a single DB is often challenging. In a blockchain-based solution, this challenge is increased multifold.

First of all, if one service provider migrates to a blockchain solution, then its operational cooperation with a traditional centralized service provider will immediately stop, as there is no default interfacing between blockchain and non-blockchain systems. The magnitude of this problem can be understood by the fact that a service provider has to migrate all of its hospitals to the blockchain system simultaneously or risk noncooperation among its own service points. Second, if all the cooperating service providers migrate to blockchain solutions, they may still face unification issues. Figure 3 shows three types of solutions in this regard. In the first unified blockchain solution, all of the e-health-care service providers connect to a single blockchain, which is maintained by either a consortium or by the government. Additionally, all eHSs maintain their independent local servers and only send trades that involve multiple eHSs. This can be viewed as a hybrid solution, which may allow some eHSs to operate a traditional system using an interface for a blockchain backbone.

The other systems shown in Figure 3 form a multiple-blockchain solution, where each service provider has an independent blockchain, which is then connected to other blockchains for interoperability. Here, either the eHS can make its whole peer network a part of the unified chain or restrict some of the peers to be part of the global chain while the others remain

local to the chain. This is a more complex solution but also allows for individual eHSs to have their own independent blockchain. Here, the solution for a traditional system to migrate to blockchain interfacing is an important design issue. In any of the aforementioned solutions, the challenges highlighted in the following sections must be addressed.

### Interoperability

This enables one eHS to exchange data with another eHS without interpreting the data. It leads to increased patient engagement and easier access, boosts efficiency, and, to some extent, enables regulatory compliance. From an engineering perspective, interoperability can be classified in one of two ways.

**Structural interoperability.** This allows for the exchange of data, and neither of the systems needs to change the format of the data, which are stored and used without any interpretation.

**Semantic interoperability.** This allows for the data to be understood by the systems without any modification. This means that not only is the structure of data the same but the data's meaning is also the same (for example, temperature stored as an integer but understood in Celsius or Fahrenheit). It is also important to note that EHR interoperability standards such as Fast Healthcare Interoperability Resources (FHIR)<sup>18</sup> are mainly implemented at the application level. The storage of EHR is usually different due to storage and query-optimization issues; however, efforts can be made to store EHRs in their native FHIR format as part of the trades, which may lead to improved interoperability.

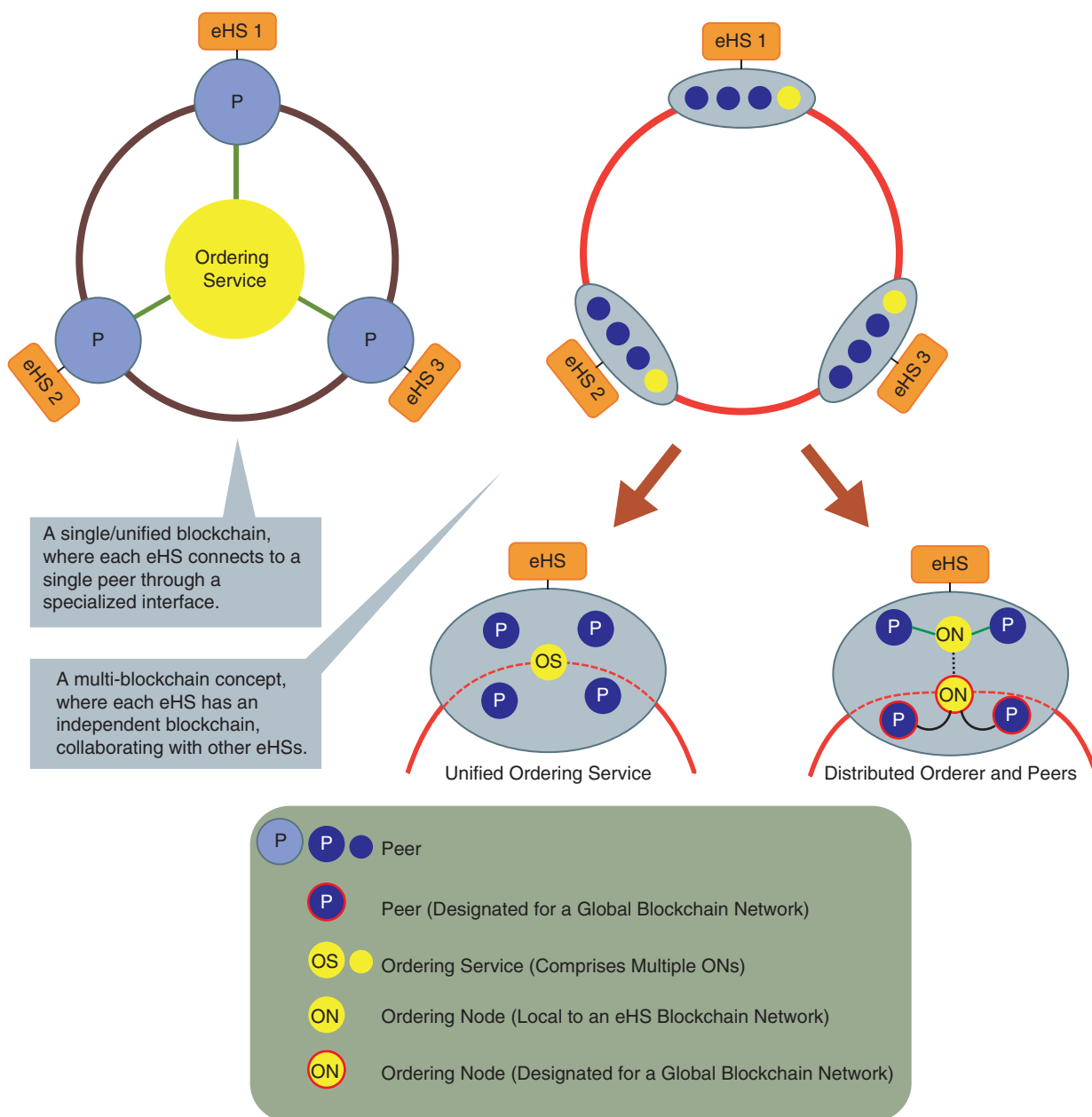
### Trade structure

Although BBCs allow unstructured data, their block and trade structures are fixed. For example, Hyperledger-based blockchain has a block header, transaction payload, and metadata as part of the block. Each component has several parameters that represent the unique information of a trade. This format may not be compatible with other protocols (for example, Ethereum). To ensure cooperation, all eHSs must be able to follow the same trade format, which is difficult. A solution to this can be in type-length-value (TLV) fields, where each part of a block is represented by a TLV. Moreover, if all of the participating systems agree on the minimum required TLVs in a block, then their order or extra TLVs will not matter. This can be an interesting research direction, as the TLV use can also enable an FHIR native format for EHR exchange among different partners.

### Storage of ledgers

Once the block is formed after consensus, the ordering service sends it to all of the peers, who add it to their ledgers and update the world state. This is the final commit process of BBC.<sup>19</sup>

The sharing of a block with all the peers is an interesting issue, especially if a multiple-blockchain solution is being used. Assume that two eHSs are involved in a trade, which is going to be part of block  $B_1$ . In a unified blockchain,  $B_1$ , after consensus formation, has to be sent to all of the peers. The number of trades generated by each eHS can be very large; hence, the memory requirements of the ledger could be astronomically high. Efficient storage solutions become an interesting research area for this problem. On the other hand, if a multiple-blockchain solution is adopted, then the



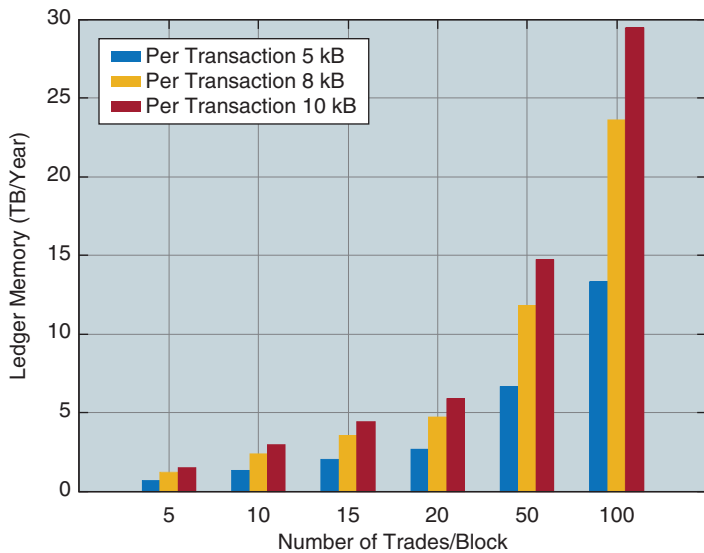
**FIGURE 3.** Interoperability solutions for cooperative eHSs.

participating eHS may choose to store  $B_1$  in its own peers only. This limits the replication of the block but may create access issues if the same patient visits

a third eHS that requires the information to be stored in  $B_1$ . Figure 4 shows the memory required by Hyperledger Fabric at every peer. This requires an

efficient trade/block discovery mechanism across different cooperating blockchains in addition to the strict access control mechanism.





**FIGURE 4.** The increase in ledger memory requirements.

### CONSENSUS FORMATION

The fundamental questions in blockchain for noncryptocurrency applications are related to trade verification and consensus formation, and perhaps this is the most misguided research area. In a noncryptosystem such as health care, the exchange of digital assets is the trade. Hence, any consensus has to be formed to enable the valid exchange of data elements.

#### What to verify?

In an eHS, several IoT devices generate data related to patients, which need to be stored and accessed by different service providers (such as doctors and nurses) as well as by third-party services (insurance agencies). Similarly, a medical test report or a prescription written by a physician is also considered a digital asset. As soon as such a digital asset is created, it has to verify its validity, authenticity, and access level. Here again, trade structure becomes an open research challenge.

Similarly, there can be many other events that require tracking, for example, administering a drug, which must be recorded as a trade. Requesting patients’ old EMRs is also a trackable event. Although this does not generate a digital asset, it is based on proper access control. Hence, a query trade must be done for this purpose.

The challenge is to identify the various types of trades that may occur in a blockchain for modern automated eHSs. In addition, the trade structure must be flexible enough for storing such dynamic information.

#### How to verify?

The process of verification begins with the SC and ends with block creation. SCs are predefined agreements between any two parties willing to participate in a trade, and they dictate the terms of the exchange.

In the given scenario, the simplest SC will exist between a temperature sensor (as an IoT device) and a patient

monitoring system (as a software entity). In a more complex scenario, a patient can create an SC to grant the physician (or group of physicians) access to his or her EMRs. Similarly, a separate SC should be created with different access privileges for other medical staff. For example, a nurse may have the right to read only part of an EMR and the prescription information, while the physician can update all of the records. The challenge here is to enable the system to efficiently generate a diverse range of SCs. Many of these can be generated using predefined templates; however, creating or changing must ensure Byzantine fault tolerance (BFT), which means that 51% of the peers must agree to it. Scalability and ease of creation are the key points in this research challenge.

The next step is that of consensus formation, which results in block formation. Ideally, it should have two parts; however, to improve the transaction rate, measured in transactions per second (TPS), some systems perform only one. The first part is to verify the validity of each trade, which means that the trade should satisfy the associated SC, must have valid signatures of all parties, and must be endorsed by the peers. In the second part, the candidate block must be verified for valid signatures and endorsed by the peers. As performing both parts is time consuming, the challenge is to have consensus-forming protocols that are highly efficient and do not compromise on the two-part process.

It is also interesting to note that Hyperledger Fabric only performs trade verification and does not require 51% of peers to vote. The participating peers can be as few as two. This significantly

improves the TPS but also compromises BFT. However, Hyperledger Fabric also allows for pluggable consensus protocols, which makes it a flexible platform. This opens the research direction of replaceable or dynamically changeable algorithms, where peers can decide which algorithms fit best for consensus for the specific types of trades or blocks. Hence, in multiple-blockchain solutions, different blockchains participating in a trade must either follow the same consensus-formation protocol (which is too restrictive) or dynamically select one (which should be interoperable). The consensus algorithm itself should be highly scalable to work at a multiple-chain level. Similarly, SCs should be acceptable across different blockchain platforms, which is a highly complex research challenge.

## USERS AND ACCESS CONTROL

In a public blockchain, especially for cryptocurrencies, all of the users are created equal. They can generate transactions or become peers for consensus formation. However, this is not the case in health-care systems, where strict access control is required. It is important to understand the difference between access to the overall system and access to the blockchain. In the former, the user may just be locally logged in through verification of a local CA or access control list (ACL), but in the latter, the user can generate trades or query the ledger. The specific challenges in this area are addressed in the next sections.

### User diversity

An e-health-care blockchain is a specialized scenario where user-type diversity is very high. This variation in type is due to access privileges.<sup>20</sup> A patient

has full access to all of his or her trades (that is, medical history), while the guardian of a minor may have limited access. This may change over time, and the system has to adapt. Similarly, one physician may have complete access to update, while a consulting physician may only have read access and the pharmacy may only be able to view prescription trades for a patient.

A blockchain solution, by default, does not address user diversification; hence, it has to be tightly coupled with the ACL of the overall system. This tight coupling is an open research area. Moreover, this coupling should be highly scalable, especially in a unified blockchain environment.

to change the access rights, a new version is created. The channel information is stored as part of the blockchain network, and as a result, it does not rely on the ACL. Furthermore, the SCs and channels have different responsibilities and should be utilized efficiently. This is still an open research challenge, however, and the solution has to be managed within the blockchain network.

## BIG DATA

In a blockchain, the storage of information or digital asset exchange is done through transactions, where all the relevant data (images and so forth) should be a part of the transaction. The

**SCs ARE PREDEFINED AGREEMENTS BETWEEN ANY TWO PARTIES WILLING TO PARTICIPATE IN A TRADE, AND THEY DICTATE THE TERMS OF THE EXCHANGE.**

### Access control and channel management

The coupling described previously is a complicated solution because it requires every trade to first be cross-checked by the ACL, which defeats the purpose of using blockchain. A better solution is to rely on SCs and channels. The concept of channels initially comes from Hyperledger Fabric, where each user/device is assigned a logical path for connecting to a peer. The solution to access control can then be implemented using these paths (channels). The channel should be bound to the patient and may have different versions. Whenever the patient wants

transactions (in the form of blocks) are stored in a file-based ledger, which cannot store large images. The typical size of a single block in any blockchain system is limited to a few megabytes, as they directly impact the performance of the system.

An eHS is heavily dependent on medical imagery (X-rays, CT scans, and so on), as discussed previously. This reason alone may make a blockchain implementation impractical in eHSs. The solution can be found in off-chain storage, but this requires several modifications in the way trades are done and data are stored. First, the off-chain storage should not allow any access other

### ABOUT THE AUTHORS

**SUJIT BISWAS** is an assistant professor with the Computer Science and Engineering Department in the Faridpur Engineering College at the University of Dhaka, Bangladesh. His research interests include the Internet of Things, blockchain, mobile computing security and privacy, big data, machine learning, and data-driven decision making. Biswas received a Ph.D. in computer science and technology from Beijing Institute of Technology, China. He is a Member of the IEEE. Contact him at [sujitedu@bit.edu.cn](mailto:sujitedu@bit.edu.cn).

**KASHIF SHARIF** is an associate research professor with the Beijing Institute of Technology, China. His research interests include wireless communication, blockchain and distributed ledger technology, programmable networks, and data-centric networks. Sharif received a Ph.D. in computing and informatics from the University of North Carolina at Charlotte. He is an associate editor of *IEEE Access*. He is a Member of the IEEE. Contact him at [kashif@bit.edu.cn](mailto:kashif@bit.edu.cn).

**FAN LI** is a professor with the Beijing Institute of Technology, China. Her research interests include wireless networks, ad hoc and sensor networks, and mobile computing. Li received a Ph.D. in computer science from the University of North Carolina at Charlotte. She is a Member of the IEEE. Contact her at [fli@bit.edu.cn](mailto:fli@bit.edu.cn).

**SARAJU P. MOHANTY** is a professor in the Department of Computer Science and Engineering, the University of North Texas, Denton. His research interest is in smart electronic systems. Mohanty received a Ph.D. in computer science and engineering from the University of South Florida, Tampa. He is the editor-in-chief of *IEEE Consumer Electronics Magazine*. He is Senior Member of the IEEE. Contact him at [saraju.mohanty@unt.edu](mailto:saraju.mohanty@unt.edu).

DB structures. This will essentially enable the use of storage and query efficiency of DB systems while securing them within the working principle of a blockchain system.

### COMPLETING THE ECOSYSTEM

Finally, EHR management is not the only process in a health-care facility. Many other departments such as accounting, human resources, pharmaceutical logistics, emergency services, and so forth are integrated into the ecosystem. As previously described, migrating one of these processes to blockchain will create a significant impact on interdepartmental communication. Most research in blockchain for health care is focused on EHRs; however, the elements and their interaction, as shown in Figure 2, are extremely important. A viable and deployable blockchain solution will only work if all entities in the ecosystem are in sync. Accordingly, the research community needs to work on blockchain and non-blockchain system interfacing while ensuring that one does not compromise the other.

than what is authorized by the peer or the ordering service. Because the objective of a DL is to have replicated, immutable copies of data, off-chain storage must ensure immutability, distributed nature, and access by verified users only. Accordingly, this demands protocol changes for query trades. Second, to add data to the off-chain storage and relate the data to a specific trade, the trade must contain a pointer to its storage location. This pointer can be

to a hash value or some other efficient mechanism within the storage. Third, the security of off-chain storage should be assured. Just as a CA is assumed to be secure and trusted, practical guarantees for storage should be ensured.

All of these requirements become design-level challenges for blockchain implementations in e-health care (and big data) scenarios. An interesting idea is to consider blockchain as a shell around existing and traditional

**T**he objectives of this article were to enable the reader to understand the complexity of implementing a blockchain solution for eHSs and to look for possible solutions. Health care is not an isolated network; hence, the blockchain solutions implemented by individual health-care service providers must be interoperable, which will require new protocols for trade and consensus management. Big data management and security in off-chain storage must be an integral part of the ecosystem. Finally, blockchain is an exciting and efficient solution

for many security- and accountability-challenged organizations; however, the migration of existing systems has a long road ahead, and the first step is to understand the needs of the application domain. Many of the design questions raised in this article may also be appropriate for other domains. ■

## ACKNOWLEDGMENTS

This work is partially supported by the National Natural Science Foundation of China, under grants 61772077 and 61370192, and by the Beijing Natural Science Foundation, under grant 4192051. Dr. Sharif and Dr. Li are cocorresponding authors.

## REFERENCES

1. H. Jin, Y. Luo, P. Li, and J. Mathew, "A review of secure and privacy-preserving medical data sharing," *IEEE Access*, vol. 7, pp. 61,656–61,669, 2019. doi: 10.1109/ACCESS.2019.2916503.
2. T. McGhin, K.-K. R. Choo, C. Z. Liu, and D. He, "Blockchain in healthcare applications: Research challenges and opportunities," *J. Netw. Comput. Appl.*, vol. 135, pp. 62–75, June 2019. doi: 10.1016/j.jnca.2019.02.027.
3. D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Senaviratne, "Blockchain for secure EHRS sharing of mobile cloud based e-health systems," *IEEE Access*, vol. 7, pp. 66,792–66,806, 2019. doi: 10.1109/ACCESS.2019.2917555.
4. W. Mougayar, *The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology*. Hoboken, NJ: Wiley, 2016.
5. P. Tasca and C. J. Tessone, "A taxonomy of blockchain technologies: Principles of identification and classification," *Ledger*, vol. 4, pp. 1–12, Feb. 2019. doi: 10.5195/ledger.2019.140.
6. S. Biswas, K. Sharif, F. Li, S. Maharjan, S. P. Mohanty, and Y. Wang, "PoBT: A lightweight consensus algorithm for scalable IoT business blockchain," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2343–2355, Mar 2020. doi: 10.1109/JIOT.2019.2958077.
7. "Hyperledger business blockchain technology," Hyperledger. Accessed on: May 15, 2020. [Online]. Available: <https://www.hyperledger.org/projects>
8. P. Sundaravadivel, E. Kougianos, S. P. Mohanty, and M. K. Ganapathiraju, "Everything you wanted to know about smart health care: Evaluating the different technologies and components of the internet of things for better health," *IEEE Consum. Electron. Mag.*, vol. 7, no. 1, pp. 18–28, Jan. 2018. doi: 10.1109/MCE.2017.2755378.
9. S. Allen, "2020 global health care outlook," Deloitte, London, 2019. [Online]. Available: <https://www2.deloitte.com/global/en/insights.html>
10. P. Pace, G. Aloï, R. Gravina, G. Caliciuri, G. Fortino, and A. Liotta, "An edge-based architecture to support efficient applications for healthcare industry 4.0," *IEEE Trans. Ind. Inform.*, vol. 15, no. 1, pp. 481–489, Jan. 2019. doi: 10.1109/TII.2018.2843169.
11. N. Kshetri, "Blockchain and electronic healthcare records [Cyber-trust]," *Computer*, vol. 51, no. 12, pp. 59–63, Dec. 2018. doi: 10.1109/MC.2018.2880021.
12. A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using blockchain for medical data access and permission management," in *Proc. Int. Conf. Open and Big Data*, Aug. 2016, pp. 25–30. doi: 10.1109/OBD.2016.11.
13. R. Jayaraman, K. Salah, and N. King, "Improving opportunities in healthcare supply chain processes via the Internet of Things and blockchain technology," *Int. J. Healthcare Inform. Syst. Informat.*, vol. 14, pp. 49–65, Feb. 2019. doi: 10.4018/IJHISI.2019040104.
14. V. P. Yanambaka, S. P. Mohanty, E. Kougianos, and D. Puthal, "PMsec: Physical unclonable function-based robust and lightweight authentication in the Internet of Medical Things," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 388–397, Aug. 2019. doi: 10.1109/TCE.2019.2926192.
15. A. D. Dwivedi, G. Srivastava, S. Dhar, and R. Singh, "A decentralized privacy-preserving healthcare blockchain for IoT," *Sensors*, vol. 19, no. 2, p. 326, 2019. doi: 10.3390/s19020326.
16. "Health insurance portability and accountability," U.S. Dept. of Human and Health Services, Washington, D.C. Accessed on: May 15, 2020. [Online]. Available: <https://hhs.gov/hipaa/>
17. General Data Protection Regulation. Accessed on: May 15, 2020. [Online]. Available: <https://gdpr-info.eu>
18. "Fast healthcare interoperability resources (FHIR) R4," Health Level Seven, Ann Arbor, MI. Accessed on: May 15, 2020. [Online]. Available: <https://www.hl7.org/fhir/>
19. S. Biswas, K. Sharif, F. Li, B. Nour, and Y. Wang, "A scalable blockchain framework for secure transactions in IoT," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4650–4659, June 2019. doi: 10.1109/JIOT.2018.2874095.
20. W. Zhang, Y. Lin, J. Wu, and T. Zhou, "Inference attack-resistant e-healthcare cloud system with fine-grained access control," *IEEE Trans. Serv. Comput.*, to be published. doi: 10.1109/TSC.2018.2790943.





# A Prize, a Prediction, and a Drama

David Alan Grier, Djanghe, LLC

*Sometimes our body of knowledge takes us to the movies and connects us to communities well outside the IEEE Computer Society.*

**W**e like to think that we are involved in dramatic and exciting work. We certainly talk a good game. Changing the world! Re-inventing society! Disrupting! Yet, anyone who has spent any time in computer science knows that it is a game of inches, not miles. Much of our work is fine and detailed. We do it hunched over a keyboard, with our faces illuminated by the light of a screen, and hope that the result will gain us 2% or 5% or maybe even 10%. This month's article from the body of knowledge both confirms our conventional model and gives us at least a suggestion of drama along with a reminder of how our contributions fit into the wider world (see "Article Facts").

Our article is "Matrix Factorization Techniques for Recommender Systems," by Yehuda Koren, Robert Bell, and Christopher Volinsky.<sup>9</sup> It ranks second on the list of the

most influential articles that have been published by *Computer*, with almost 14,000 downloads and more than 2,000 citations. (Articles for the "Body of Knowledge" column are taken from a report prepared by the IEEE publications office on 15 November 2019, and the statistics were current as of that date. Other citation

services can and do give different numbers.) At its base, it reminds us of several fundamental lessons of computer science. It illustrates the growing importance of data analytic methods to computing. It shows how computing borrows

## ARTICLE FACTS

- » Article: "Matrix Factorization Techniques for Recommender Systems"
- » Authors: Yehuda Koren, Robert Bell, and Christopher Volinsky
- » Citation: *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009
- » *Computer* influence rank: #2 with 13,872 downloads and 2,001 citations



and adapts ideas from other fields. Finally, it truly tells that tale of progress by inches.

The article shows how to apply matrix-factoring methods, a topic originally explored by the French mathematician during the 1870s, to recommender systems, a type of computer system that was fairly new in 2009. Recommender systems are programs that suggest products for customers to buy. They expand the role of the old store clerk, who could tell a customer that “since you have selected that one product, you might want to consider also purchasing this one.” They have grown rapidly in importance with the rise of Internet commerce, though they actually predate that increase by a few years. One of the first articles to employ that term is a 1997 survey that looks at network programs that recommend news, websites, or people. It describes a group of systems that is long gone, but it presents a set of concepts that remains part of the field.<sup>6</sup>

In the modern version of these systems, vectors contain the data for both people and products. Matrices map the one onto the other, people and their tastes onto products and their features. Matrix factorization reveals the structure of this mapping: how personal propensities determine which goods will be sold. This article shows how to factor matrices in an efficient way and interpret their components. Efficiency is important. The data sets can be large and are certainly dynamic. A popular product this year may vanish before the next. The needs of a customer with a newborn baby are different from one with a five-year-old. This observation leads us to the drama of the story.

Our article, “Matrix Factorization Techniques for Recommender Systems,” was not done as an abstract bit of research in an academic lab. It was one of the last articles to come from an innovation

contest, which was run by Netflix. In October 2006, Netflix wanted to improve its recommendation system. At the time, the company rented videos using the technology of DVDs and the communications system of the U.S. Postal Service. It would send DVDs to customers, who would view the movies and return them to Netflix. Customers would then use a website to rate their rental and choose their next video, using a recommender system called CinemaTech.

Netflix wanted to reduce the errors in its recommendation system, the number of times that a recommended DVD was rejected by a customer. So, the company released a data set with 100 million transactions by 480,000 customers regarding 17,770 movies and challenged any person or organization to create a recommender system that was 10% better (as measured by a decrease in the root-mean-square error). It offered a prize of US\$1 million.<sup>1</sup>

This contest was an example of crowdsourcing, a series of techniques that was growing rapidly during the first decade of the 21st century. Behind crowdsourcing was the idea that you could market mechanisms to manage work, even the work of generating new ideas. Such markets have been common in many parts of the economy. The design and architecture community has relied on them for decades to evaluate ideas. During the early days of the aeronautical industry, different organizations used them to improve the technology of airplanes. Within the field of computer science, we saw them used by DARPA to improve the technology for self-driving cars. Finally, we have seen a number of companies build successful business models on crowdsourcing.

When using crowdsourcing, an organization needs to decide how much it will manage the process. In theory, one should be able to put a problem to the crowd and then wait for the

multitude to provide one with a solution. For complicated problems, such as the Netflix challenge, things rarely work that way. We can often look at a crowd and identify the expertise that will solve the problem, but we will quickly realize that one person or team does not possess all the expertise that is needed to produce a workable solution. As a result, we have to manage the crowd. We have to push teams together, combine promising ideas, and suggest profitable directions for work.<sup>8</sup>

In the Netflix challenge, the authors of our article quickly produced a system with a substantial improvement of 7.42%.<sup>3</sup> However, they were unable to get their system above the 10% target. Other groups were able to get similar improvements and claim the lead, at least for a time. No one was able to clear the 10% mark. As the context progressed, Netflix pushed and prodded groups together. One of the early outcomes of the contest was to reveal expertise that research groups had not known. In particular, they learned of a group in Hungary that had some good ideas.<sup>3</sup> In June 2009, a combined group that included our authors reached the 10% goal and claimed the prize.<sup>4,7</sup>

While this contribution to the body of knowledge came out of a contest, it would be wrong to view it as coming from outside the computer science community. The knowledge discovery and data-mining community embraced the challenge. It organized a workshop on the Netflix challenge as part of its annual conference. Many were intrigued by the data set alone. “Until now,” one organizer noted, “researchers who have been working to improve recommendation systems have been relying on a much smaller database.”<sup>2</sup>

Unlike many a current drama, the Netflix challenge had no sequel. Though the company organized a second competition, it quickly ended the contest when it learned that its data set was

not as private as it had thought. A pair of researchers showed how the test data could be used to identify customers and their preferences.<sup>5</sup> So, after a stellar moment in the public eye, researchers in recommender systems returned to the private work of inches and percentages. But to do that work, those researchers had the results that were published in this article. **[C]**

## REFERENCES

1. R. M. Bell and Y. Koren, "Lessons from the Netflix Prize Challenge," *SKDD Explor.*, vol. 9, no. 2, 2008.
2. K. Hafner, "And if you liked the movie, a Netflix contest may reward you handsomely," *NY Times*, Oct. 2, 2006. [Online]. Available: <https://www.nytimes.com/2006/10/02/technology/02netflix.html>
3. K. Haffner, "Netflix Prize still awaits a movie seer," *NY Times*, June 4, 2007. [Online]. Available: <https://www.nytimes.com/2007/06/04/technology/04netflix.html>
4. K. Haffner, "Netflix Prize: Close, but no \$1 million cigar," *NY Times*, Nov. 13, 2007.
5. A. Narayanan and V. Shmatikov, "Robust de-anonymization of large datasets (how to break anonymity of the Netflix Prize dataset)," in *Proc. IEEE Symp. Security and Privacy*, Feb. 5, 2008.
6. P. Resnik and P. Varian, "Recommender systems," *Commun. ACM*, vol. 40, no. 3, pp. 56–58, Mar. 1997. doi: 10.1145/245108.245121.
7. R. Rocha, "Montreal team shares lead: Netflix Prize," *The Gazette*, July 7, 2009. [Online]. Available: <https://www.pressreader.com/canada/montreal-gazette/20090707/textview>
8. J. A. Villarreal, J. E. Taylor, and C. L. Tucci, "Innovation and learning performance implications of free revealing and knowledge brokering in competing communities: Insights from the Netflix Prize challenge," *Comput. Math. Organ. Theory*, vol. 19, pp. 42–77, Mar. 2013. doi: 10.1007/s10588-012-9137-7.
9. Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009. doi: 10.1109/MC.2009.263.

**DAVID ALAN GRIER** is a principal with Djaqhe, LLC. He is a Fellow of the IEEE. Contact him at [grier@gwu.edu](mailto:grier@gwu.edu).



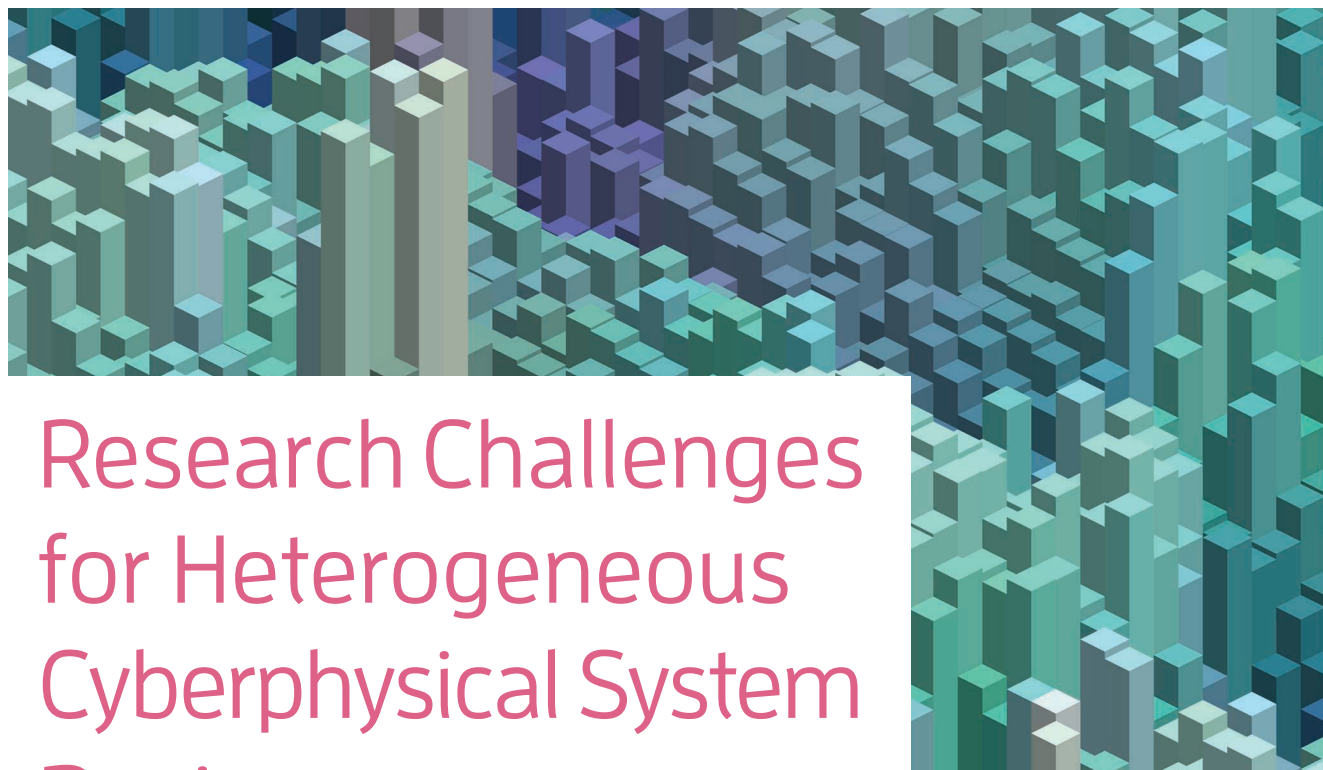
**IEEE Security & Privacy magazine provides articles with both a practical and research bent by the top thinkers in the field.**

- 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](http://computer.org/security)





# Research Challenges for Heterogeneous Cyberphysical System Design

**Shuvra S. Bhattacharyya**, University of Maryland

**Marilyn C. Wolf**, University of Nebraska

*Heterogeneous computing is widely used at all levels of computing. However, heterogeneity presents challenges. This article considers research issues in heterogeneous cyberphysical system design, including interoperability, physical modeling, models of computation, self-awareness and adaptation, architecture, and scheduling.*

**H**eterogeneous computing offers the potential to streamline the execution of key tasks for processing, sensing, actuation, and communication using devices that are better suited to those applications than architectures composed of collections of identical mechanisms. This potential is of great utility for cyberphysical systems (CPSs), where constraints on energy consumption, cost, and real-time performance often motivate the investigation of highly streamlined solutions. However, increased use of heterogeneity leads to complex challenges and important needs associated with interoperability and model-based design in CPSs. This column outlines challenges in heterogeneous CPS design and motivates the need for approaches to system-level architecture that are based on complementary collections of compact system-level models.



## COMPACT SYSTEM-LEVEL MODELS

Raising the level of abstraction in design processes for CPSs can facilitate interoperability by making it easier to reason about the behavior of subsystems in an architecture and the interactions between them. However, due to the multifaceted nature of CPS design, no single abstraction or small set of abstractions is adequate for all systems. Instead, the ab-

models based on dynamic changes in the data acquired from instrumentation and, conversely, for control of the instrumentation processes by the executing models.

These motivations for diverse and compact abstractions lead us to advocate the concept of compact system-level models as a central element in the design and implementation of CPSs. Many different types of models are relevant to CPS design.

Modern CPSs involve hundreds of thousands to tens of millions of lines of high-level language code or more.

stractions to employ must be selected and applied in complementary ways that are well matched to the targeted class of applications and the objectives and constraints involved in their design.

Given the complexity of modern CPSs, the size of the models in the employed abstractions is an important consideration in their formulation and selection. The transition from assembly language to high-level languages, such as FORTRAN and C, which began many decades ago, can be considered as an increase in the level of abstraction. However, modern CPSs involve hundreds of thousands to tens of millions of lines of high-level language code or more. The compactness of the models involved in the abstractions becomes an important concern to facilitate human understanding and tractable analysis of the representations.

The strategic application of compact models is important, for example, in the paradigm of dynamic, data-driven applications systems (DDASs), where an executing model of an application is integrated into a feedback loop with instrumentation processes that supply data to the representation.<sup>1</sup> Accurate, compact facsimiles are useful for the real-time adaptation of DDAS

Some prominent examples include the following:

- *Models of physical phenomena:*<sup>2</sup> Computing is a physical act: it takes time and energy, and the reliability of the result depends on the physics of the computing system. Taking all these physical phenomena into account in multibillion-transistor systems is extremely challenging.
- *Models of computation:* A model of computation defines how an interconnected set of components interacts to perform computation. Some important classes of computation models include dataflow models, state machines, and discrete-event models. Models of computation may impose restrictions on how components are defined or interact that make important analysis or optimization problems become tractable (for example, see Eker et al.<sup>3</sup>). In contrast, fundamental analysis problems, such as whether a program halts or has bounded memory requirements, are undecidable in conventional programming languages for general-purpose computing. Computation

models contribute to modeling compactness by abstracting implementation details of functional components and their coordination.

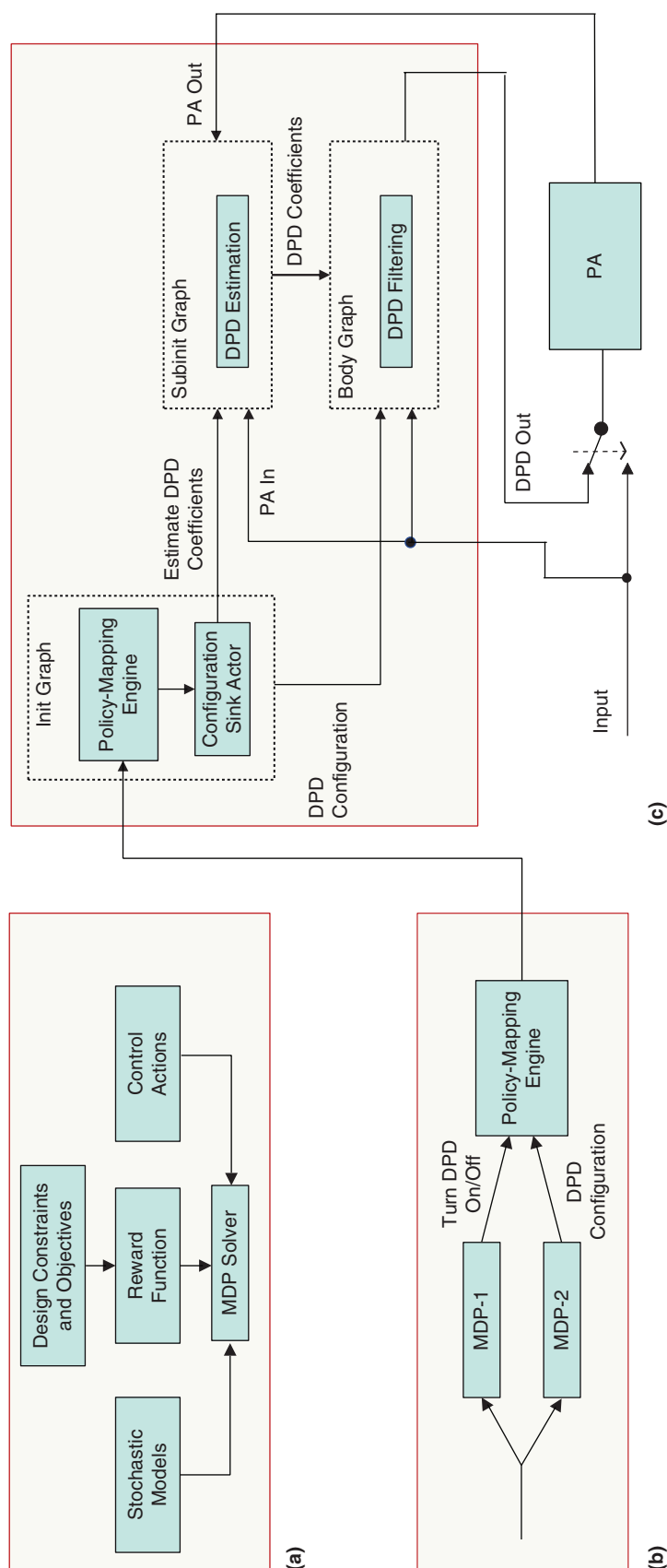
- *Models of self-awareness and adaptation:*<sup>4</sup> Stochastic models provide systems with compact, runtime-ready models that the systems can use to estimate their own state. Training enables us to capture complex models as long as we have sufficient data. Once trained, those models can be evaluated much more efficiently on the platform. Their results enable the system to reflect on its own power and thermal behavior. Managing power and thermal behavior is critical to maintaining system longevity.
- *Models of architecture:* While models of computation focus on capturing the algorithmic behavior of application systems, models of architecture provide compact abstractions of the hardware on which the algorithms are mapped.<sup>5</sup> Architecture models are formulated to enable efficient, reproducible estimation of nonfunctional costs associated with executing applications that are described in terms of a given model of computation. These costs include important metrics for efficiency evaluation, such as latency, throughput, memory requirements, and energy consumption. A key concept in the formulation of models of architecture is the decomposition of application execution into quantized units of communication and computation and the estimation of costs in terms of these abstract units. Architecture models are more constrained and operate at a higher level of abstraction compared to hardware description languages, such as Verilog and VHDL.

- › *Scheduling models*: Scheduling is an important aspect of implementation that is abstracted away by models of computation. Scheduling involves assigning computational tasks to processing resources and ordering tasks that share common resources. Scheduling often has a major impact on metrics for efficiency evaluation, including the ones listed previously. Model-based scheduling representations provide formal, platform-independent approaches for representing, reasoning about, and transforming schedules.<sup>6</sup>

A design methodology based on compact system-level models for CPSs involves the selection of such facsimiles and the definition of how representations and design tools associated with them are cooperatively applied in system design processes. While there are tradeoffs between model complexity and accuracy that may be involved in the models that are employed, restricting attention to only the highest-fidelity representations may severely limit the extent of the design space that can be investigated.

## MODELING EXAMPLE


An example of a complex subsystem design using multiple forms of compact system-level models is the Markov decision process (MDP) framework for adaptive digital predistortion (DPD) systems (MADSs).<sup>7</sup> DPD is a type of algorithm that is used to counteract nonlinearities in power amplifiers (PAs) to improve the quality of wireless communications signals. The design and configuration of DPD systems involve complex tradeoffs among signal quality, energy efficiency, and real-time performance. The MADS framework is demonstrated by mapping it into an optimized implementation on a CPU/graphics processing unit (GPU) platform. The model-based design of the MADS framework is illustrated in Figure 1.



**FIGURE 1.** The MADS framework (adapted from Li et al.<sup>7</sup>): the (a) policy generation and (b) hierarchical MDP subsystem. (c) The parameterized dataflow model.

The MADS framework illustrates an approach to several of the challenges associated with heterogeneous CPS design discussed in this column. MADS applies a model of the physics involved in a communications transmitter to define, simulate, and fine-tune the core predistortion algorithm that is employed. An MDP is employed in MADS as a model that provides self-awareness and adaptation capabilities. MDPs are probabilistic models

the dynamic manipulation of parameters in the body graph. The init and subunit graphs differ in the frequency with which the associated parameter adaptation operations are carried out, with subunit graph processes being more frequent.<sup>9</sup> In MADS, the parameterized dataflow model is used as a starting point to map the MDP-equipped adaptive system into a CPU/GPU implementation. For more details on the MADS framework, including

by such models are all representative directions for future research that can help to address the complexities and opportunities presented by heterogeneous CPS design. 

The study of design methodologies based on cooperating compact system-level modeling approaches is a broad area ripe for further study.

used to derive adaptation policies in uncertain environments. In particular, MDPs are used in the context of environments characterized using memoryless probability distributions; that is, the distribution of the next state is dependent only on the current state, not the trajectory of prior states that led to the current one. In MADS, MDP-based DPD architecture adaptation is performed with the objective of jointly optimizing signal quality, system throughput, and power consumption. In general, MDP models can become large and unwieldy to employ in complex applications. To help ensure the compactness of the MDP model that is employed, a hierarchical MDP<sup>8</sup> structure is designed, as illustrated in Figure 1(b).

Parameterized dataflow<sup>9</sup> is used in MADS as a model of computation to represent the algorithms employed for adaptation and DPD operation and characterize their interactions. In parameterized dataflow, the design for a signal processing system is decomposed into three cooperating dataflow graphs, called the init graph, subunit graph, and body graph (see Figure 1). The body graph represents the core signal processing functionality, while the init and subunit graphs describe the functionality for

the different design components illustrated in Figure 1, we refer the reader to the presentation by Li et. al.<sup>7</sup>

Many of the state-of-the-art methods for CPS design and implementation are not model based or involve a focus on individual representation types, for example, the development of software synthesis techniques for specific models of computation or reconfigurable architectures based on specific models for self-awareness and adaptivity. The study of design methodologies based on cooperating compact system-level modeling approaches is a broad area ripe for further study. For example, a deeper understanding is needed for many modeling techniques, concerning how these models may be adapted or parameterized to provide more flexible tradeoffs between compactness and accuracy. Some compact modeling adaptations, such as hierarchical and factored MDPs,<sup>8,10</sup> are established in the literature but not applied in practice to their full potential. More diverse families of compact models, increasingly sophisticated design tool support for applying and integrating them, and additional concrete ways to assess the novel tradeoffs introduced

## ACKNOWLEDGMENTS

We thank the following people who contributed to discussions about heterogeneous computing and interoperability that have influenced this article: Alvaro Cardenas, Roger Chamberlain, Tam Chantem, Changhee Jung, Miriam Leiser, Shivakant Mishra, Mahdi Nikdast, Massimiliano Pierobon, Aviral Shrivastava, Heechul Yun, and Ting Zhu. This work was supported in part by the U.S. National Science Foundation under grants CNS1514425 and CNS151304 and the U.S. Air Force Office of Scientific Research under grant FA9550-18-1-0068.

## REFERENCES

1. E. P. Blasch, S. Ravela, and A. J. Aved, Eds., *Handbook of Dynamic Data Driven Applications Systems*. New York: Springer-Verlag, 2018.
2. M. Wolf, *The Physics of Computing*. San Mateo, CA: Morgan Kaufmann, 2016.
3. J. Eker et al., "Taming heterogeneity: The Ptolemy approach," *Proc. IEEE*, vol. 91, no. 1, pp. 127-144, Jan. 2003. doi: 10.1109/JPROC.2002.805829.
4. N. Dutt, A. Jantsch, and S. Sarma, "Toward smart embedded systems: A self-aware system-on-chip (SoC) perspective," *ACM Trans. Embedded Comput. Syst.*, vol. 15, no. 2, p. 27, Feb. 2016. doi: 10.1145/2872936.
5. M. Pelcat et al., "Reproducible evaluation of system efficiency with a model of architecture: From theory to practice," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 10, pp. 2050-2063, Oct. 2018. doi: 10.1109/TCAD.2017.2774822.
6. S. S. Bhattacharyya and J. Lilius, "Model-based representations for dataflow schedules," in *Principles of Modeling: Essays Dedicated to Edward*

- A. Lee on the Occasion of His 60th Birthday, M. Lohstroh, P. Derler, and M. Sirjani, Eds. New York: Springer-Verlag, 2018, pp. 88–105.
7. L. Li et al., “MADS: A framework for design and implementation of adaptive digital predistortion systems,” *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 4, pp. 712–722, Dec. 2019. doi: 10.1109/JETCAS.2019.2952145.
  8. A. Jonsson and A. Barto, “Causal graph based decomposition of factored MDPs,” *J. Mach. Learn. Res.*, vol. 7, pp. 2259–2301, Dec. 2006. doi: 10.5555/1248547.1248628.
  9. B. Bhattacharyya and S. S. Bhattacharyya, “Parameterized dataflow modeling for DSP systems,” *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2408–2421, Oct. 2001. doi: 10.1109/78.950795.
  10. C. Boutilier, R. Dearden, and M. Goldszmidt, “Exploiting structure in policy construction,” in *Proc. Int. Joint Conf. Artificial Intelligence*, 1995, pp. 1104–1111. doi: 10.5555/1643031.1643043.

**SHUVRA S. BHATTACHARYYA** is a professor at the University of Maryland, College Park, and part-time international research chair at INSA/IETR/INRIA, Rennes, France. He is a Fellow of the IEEE and senior member of the ACM. Contact him at ssb@umd.edu.

**MARILYN C. WOLF** is Koch Professor of Engineering and chair of computer science and engineering at the University of Nebraska–Lincoln. She is a Fellow of the IEEE and ACM. Contact her at mwolf@unl.edu.



# Call for Articles

## IEEE Pervasive Computing

seeks accessible, useful papers on the latest peer-reviewed developments in pervasive, mobile, and ubiquitous computing. Topics include hardware technology, software infrastructure, real-world sensing and interaction, human-computer interaction, and systems considerations, including deployment, scalability, security, and privacy.

**Author guidelines:**  
[www.computer.org/mc/pervasive/author.htm](http://www.computer.org/mc/pervasive/author.htm)

**Further details:**  
[pervasive@computer.org](mailto:pervasive@computer.org)  
[www.computer.org/pervasive](http://www.computer.org/pervasive)

**IEEE pervasive COMPUTING**  
 MOBILE AND UBIQUITOUS SYSTEMS



# Contactless U: Higher Education in the Postcoronavirus World

**Phil Laplante**, Pennsylvania State University

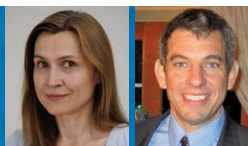
*The profound impact of the coronavirus COVID-19 has tragically affected the world, including undergraduate and graduate university education. These impacts will be very long lasting, significantly changing the nature of higher education forever. A contemporary account of what is happening and predictions on how this situation will transform higher education are offered.*

**B**y late February 2020, the world was reeling from the onslaught of COVID-19, many had already died, and the predicted trajectory was catastrophic. Worldwide, a large majority of

university administrators realized that they would need to prohibit or severely limit physical access to campuses, thereby threatening completion of the current academic session, research programs, and the conduct of university business. This dire situation could only be mitigated by virtual meetings and research activities and by reinventing courses already in session or about to launch into online equivalents. Across the globe, decisions were made to put courses online and move all meetings to virtual formats.

Faculty and students who never tried online learning before (and were reluctant) were suddenly faced with it. The implementation issues and second-order effects impacted

housing, scholarships and other funding sources, travel arrangements, and more. The forced conversion made sensational news with everyone, even outside academia, discussing the subject in email threads, on social media, and on radio and television. Teaching and taking online courses are difficult, and this fact surprised many except those with experience.



## TECHNOLOGY CHALLENGES

There are always significant technological challenges to delivering content online, but these were intensified as faculty rushed to convert courses. Lectures were hastily recorded, and there were issues with quality, high-speed network access for streaming, and compatibility. Many course materials needed to be put online for the first time, sometimes into learning management systems, or sometimes without the benefit of such systems. Simulation software for laboratory courses needed to be identified, acquired, and installed and lab activities modified accordingly.

Where synchronous communications were used, there were numerous compatibility, processing power, and configuration issues. Many university networks and cloud services were not ready for or capable of the tremendous increase in demand. On the student end, throughput and bandwidth were often not sufficient.

Many universities found their technical support departments overwhelmed or underprepared, and many professors who may have been technologically challenged were forced to face their weaknesses. During course delivery, the technological disparity in ability and access to technology for certain students were exposed.

### Impact

Universities need to address the technical weaknesses in faculty and staff, reevaluate their technical support departments, and, more significantly, address students' technological disparity.

## PEDAGOGICAL SOUNDNESS

Some faculty and administrators had already questioned the pedagogical soundness of online courses, but the COVID-19 pandemic crisis brought this issue to the fore. While lectures and advising could be delivered online, "show-stopping" challenges such as how to conduct virtual labs in chemistry,

physics, and engineering were solved, even if suboptimally. There was the usual time-zone challenge for synchronous courses and meetings. Many new problems related to internships, clinicals, and field work, both logistical and bureaucratic, needed to be resolved through telepresence or other means. In solving these problems, new issues involving mobility of students across borders in the post-COVID-19 world were identified, for example, how to handle certain programs and visa rules that don't allow for online components.

### Impact

Universities will need to provide more training and incentives for faculty on how to build and deliver pedagogically sound online courses.

## AUTHENTIC PRESENCE AND ENGAGEMENT

An important aspect of online courses is providing students with a sense that they are having ample interaction with the professor and other students and not just self-teaching, that is, providing an "authentic presence." Creating an authentic presence, whether with live streaming, asynchronous learning, directed activities, or a combination, is difficult. Engagement involves creating a "community" of learners, and this can also be difficult to do in virtual communities. Faculty teaching online for the first time, and even those with experience, had to provide engagement often in a hastily constructed "classroom" with scared or unenthusiastic

students. Using appropriate online learning management systems and providing supporting media and activities have already been part of the solution, but professors and students explored innovative ways of engagement during the COVID-19 crisis.

For teleconferenced meetings, attendees had to worry about what they looked like (showered, shaven, made up), what to wear (business professional, casual, pajamas), where to set up their attendance (appropriate venue or sitting on a bed), and how to organize

Universities need to address the technical weaknesses in faculty and staff, reevaluate their technical support departments, and, more significantly, address students' technological disparity.

the background. Many online attendees worried about stifling academic freedom online. For example, professors and students were less likely to say controversial things if they are written in text messages or clearly being recorded as part of the course (as opposed to surreptitiously and possibly illegally being recorded, as now occurs).

University administrators also confronted the restrictions of employer tuition reimbursement and stipend restrictions for online courses. In many cases, reimbursement is lower for online courses; this needed to be ignored during the pandemic. In addition, onsite attendance requirements for courses, presentations, internships, dissertation defenses, and so forth had to be relaxed.

### Impact

Universities and employers may have to reconsider relaxing certain residency restrictions permanently.

### ASSESSMENT AND GRADING

Assessing student learning and assigning a grade are major challenges for online courses. While many residential courses already use online testing (at testing centers or over secured remote systems), some question the efficacy of this mode. Many professors find that it might be easier for students to cheat when not under the watchful eye of the professor or a graduate assistant. Proctoring of re-

### SECURITY AND PRIVACY

Some of the most serious objections to online education involve security and privacy. Unfortunately, the COVID-19 pandemic highlighted the validity of these concerns.

Throughout the pandemic, news outlets reported teleconferencing security issues, phishing attacks, and network vulnerabilities. In teleconferencing, services like Zoom were exploited, potentially allowing unauthorized users

world is operating virtually, then why not go to school virtually (and work part time)?” Similarly, the economic damage done by the COVID-19 pandemic will require students to defer college and choose online education, thus saving on room and board and allowing them to work full or part time at home (or help at home) while getting their education.

During the COVID-19 pandemic, universities held all meetings via teleconferencing. Those who were previously reluctant to hold or attend such meetings may now be less likely to object to this format, since they are easier to attend (and greener since they cut down on driving to campus). Virtual meetings may have improved faculty and other stakeholder engagement and removed many excuses for participation.

Jobs that were previously thought to be “in residence only” were performed partially or entirely at a distance, challenging previously held notions, and perhaps changing the way things are done on campus forever. Administrators who objected to the work level performed remotely discovered that some people may work harder when given the flexibility of working from home.

#### Impact

Administrators, professors, and students have gained confidence to work in remote environments, and university administrators may be convinced that online learning can significantly lower the cost for students and the university. Universities will have to face these realities and bring more courses and programs online. Universities that already had a strong online presence will have a huge advantage.

### PREDICTIONS

There are many advantages and disadvantages to online education, and the COVID-19 pandemic portrayed these in sharp relief. Online learning is not for everyone nor practical for some courses, but its role in the higher education

Universities that already had a strong online presence will have a huge advantage.

ote exams is often done via telepresence by the university or a third party, but during the COVID-19 pandemic, third-party monitoring services such as Examity were overwhelmed. One way around these problems is to offer open-book exams, but identity management is still an issue. In addition, there have always been issues with online learning involving securing exams and online material, stolen curricula, and overall security to learning management systems. The COVID-19 pandemic, once again, reminded us of these problems.

During the crisis, many universities adopted relaxed grading, that is, students could elect to earn a pass/fail grade only. This approach is thought to reduce cheating on remote tests, relieving the pressure of achieving a certain grade. It also takes into consideration that students working from home may have unique obstacles, for example, taking care of children or having to share computers at home, all likely scenarios during the pandemic. However, pass/fail grading presents other problems with respect to tuition reimbursement policies and minimum GPA requirements.

#### Impact

Assessment and grading policies and practices for online courses will need to be reevaluated.

to attend (bomb) meetings, potentially disrupting classes, dissertation defenses, and university business. Phishing attacks included a variety of official-looking university emails that directed recipients to malicious sites, caused them to reveal confidential information, or contained harmful payloads. Malware may have been resident in some of the freeware software applications hastily selected for courses.

#### Impact

New attention to security and privacy issues, both old and new, is needed. Universities must invest more resources into fixing these problems.

### CHANGING ATTITUDES

Many university officials and professors have seen their objections to the remote delivery of courses disappear, as out of necessity courses were flipped on very short notice from residential to remote format. More profoundly, and surprisingly, many parents, students, and tuition sponsors (for example, employers and government grant agencies) have begun to question their objections to online education and even the need for a residential component. These stakeholders are now asking questions such as “If students can complete their programs virtually from home, why do they need to stay on campus?” and “If the business

mix will exponentially increase. The COVID-19 pandemic has converted many naysayers, shown how previously insurmountable problems could be overcome, exposed waste, and spurred many innovations for online learning. Similarly, the way that academia conducts business has been transformed with the mandate for remote work. In the post-COVID-19 world, higher education will never be the same. Here are some of the ways I predict it will change.

- › Many students will opt to take some time off to let the COVID-19 pandemic subside, but this number can be mitigated by online alternatives.
- › Of those who choose to return to school, many will be reluctant

to attend any in-person classes or at least close-contact courses (such as labs and large seminars) for the near future, if not forever, thus driving up enrollments in online learning.

- › Socioeconomic and political forces will result in fewer international students, driven by an abundance of caution and travel, funding, and visa restrictions.
- › Many schools will lower admissions standards for residential students due to decreased admissions applications.
- › The justification for certain meetings, infrastructure, hardware, and equipment on site or at all will be strongly challenged.

- › Academia will use new ways of working to include significantly more online interaction.

During the COVID-19 pandemic, restaurants advertised “contactless delivery.” Will we see the same promoted in higher education? As computer science professionals, we can act as resources and advocates to those who challenge the efficacy of online learning, are afraid to try it, or just don’t know how to make it work. ■

**PHIL LAPLANTE** is a professor at Pennsylvania State University, University Park. Contact him at [plaplante@psu.edu](mailto:plaplante@psu.edu).

## IT Professional

TECHNOLOGY SOLUTIONS FOR THE ENTERPRISE

# CALL FOR ARTICLES

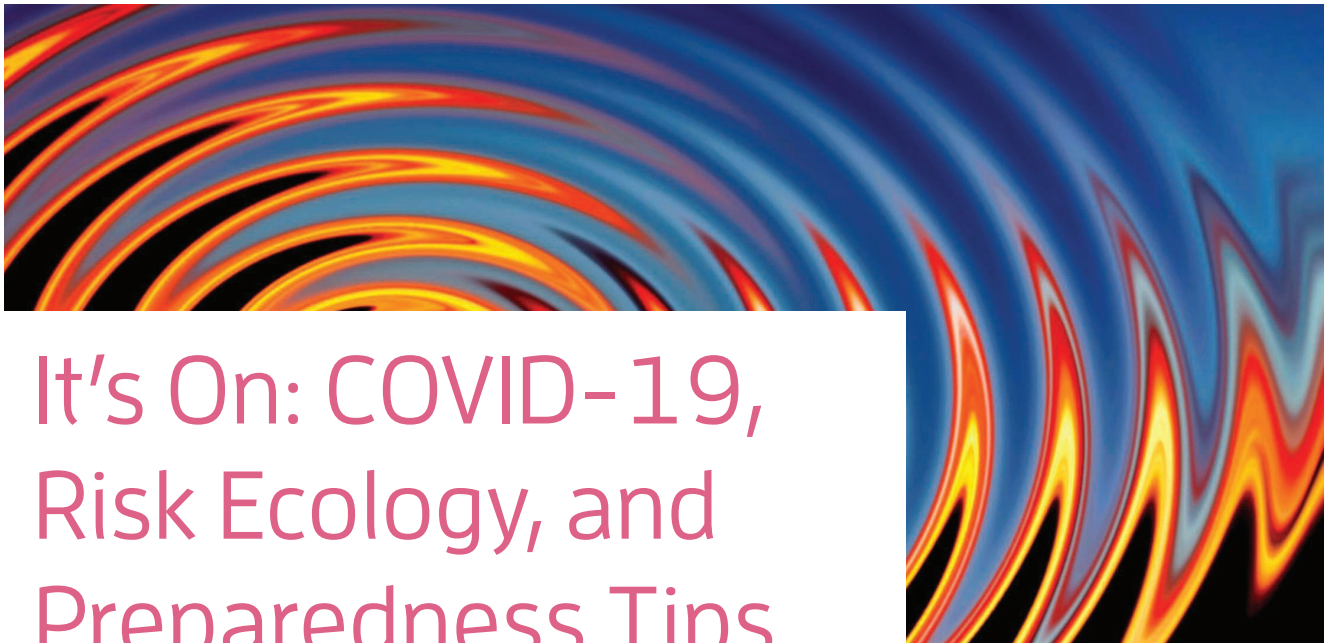
*IT Professional* seeks original submissions on technology solutions for the enterprise. Topics include

- emerging technologies,
- cloud computing,
- Web 2.0 and services,
- cybersecurity,
- mobile computing,
- green IT,
- RFID,
- social software,
- data management and mining,
- systems integration,
- communication networks,
- datacenter operations,
- IT asset management, and
- health information technology.

We welcome articles accompanied by web-based demos. For more information, see our author guidelines at [www.computer.org/itpro/author.htm](http://www.computer.org/itpro/author.htm).

## WWW.COMPUTER.ORG/ITPRO





# It's On: COVID-19, Risk Ecology, and Preparedness Tips

**Hal Berghel**, University of Nevada, Las Vegas

**Robert N. Charette**, ITABHI Corporation

**Edward G. Happ and John Leslie King**, University of Michigan

*"It's on" can mean something big is happening, like the COVID-19 pandemic. In slang terms, it's always on: disruptions are going to happen. This article explores emergency crises, preparedness, risks, surprise, risk fatigue, and tips.*

**A**s this issue goes to press, "it's on" can refer to the disruptive COVID-19 pandemic. Millions of people will be infected and hundreds of thousands will die before the pandemic ends. "Social distancing" has entered the lexicon. Companies and agencies have shut down. Economic losses are stratospheric. However, in slang terms, it's always on: disruptions are going to happen, including pandemics,

earthquakes, tsunamis, floods, droughts, and wars. Yet, after the fact, people complain that they were unprepared. This article explores emergency crises, preparedness, risks, surprise, risk fatigue, and tips.

The authors have more than a combined 160 years of experience with computing theory, research, design, construction, application, and evaluation. In 1958, March and Simon said scheduled work drives out unscheduled work.<sup>1</sup> Crises drive out everything. Dependence on computerized systems is growing, and the consequences of disruption multi-

ply accordingly. People in computing must do better with what former U.S. Secretary of Defense Donald Rumsfeld called "unknowns," both the known unknowns and the unknowns that often come as a surprise.<sup>2</sup>

## EMERGENCY CRISES AND PREPAREDNESS

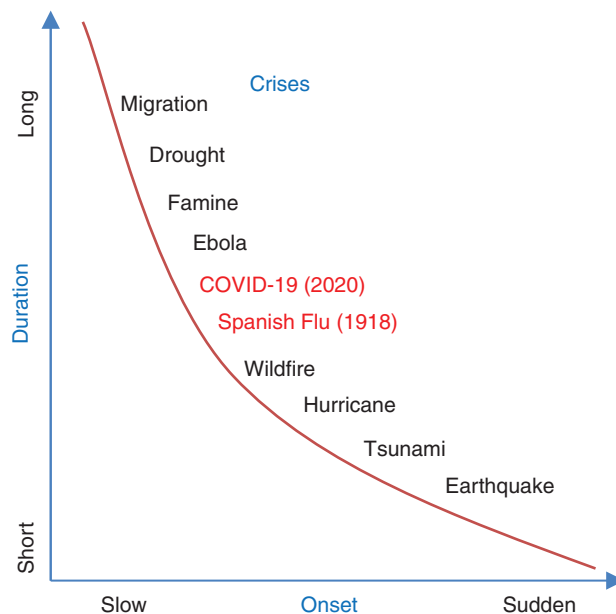
An emergency crisis requires a response: determine what is known, assess the resources available, and decide what must be done. Disasters happen all of the time and are becoming more frequent. Annual human-caused crises such as wars, terrorist attacks, or refugee migration peaked at around 250 per year in 2005 and have been declining since (although



they might tick up again).<sup>3</sup> Fires and epidemics can be either human or nature caused. Nature-caused crises such as hurricanes, typhoons, winter storms, tsunamis, earthquakes, avalanches, floods, and heat waves produce more monetary damage, sometimes costing billions of dollars, and could accelerate with climate change. Crises vary by onset, from rapid earthquakes to slow droughts, and by duration, from short like a tsunami to long like refugee immigration (see Figure 1). The consequences of crises are growing.

The cost ratio of crisis response to crisis preparedness is about 6:1.<sup>4,5</sup> However, it is difficult to know what investment in preparedness ought to be, given that not all crises have the same priority. It is challenging to predict when or where a crisis will occur or what its consequences will be. In some organizations, a 100-year event is not an executive concern, although it is in others. What is a 100-year event, anyway? The frequency of these incidents might be increasing. The oft-referenced precursor to the COVID-19 pandemic, the so-called Spanish Flu of 1918–1919, was 100 years ago. Crises can damage vulnerable things<sup>6,7</sup> and can make them “old.”

Preparedness should be as easy as spotting the intersection of crisis likelihood and cost curves, but implementation can be political. The COVID-19 pandemic has cost the U.S. government more than US\$3 trillion so far. At 6:1, preparedness would have cost US\$500 billion, nearly the annual U.S. defense budget. This money would probably not have been made available before the pandemic. Crisis preparedness might be popular



**FIGURE 1.** The onset versus duration for select natural crises.

right after the crisis, but its acceptance fades.

There are many unknowns. First responders are 90% locals, often bystanders, and it is impossible to tell in advance who they will be.<sup>8</sup> Hired responders can take too long to get to their destinations. Simple single points of failure might be covered in the backup plan (electricity, telecom, and so on), but staff to execute the plan might not be considered. Planning saves lives. The World Trade Center organizations that lost people on 9/11 had plans, although organizations without plans also lost people. With pandemics like COVID-19, first responders are vulnerable, and even cloud-based systems can be as vulnerable as premises-based ones. There are too many threats to be ready for all of them, but some can be addressed. The best is the enemy of the good: overly ambitious plans fail at run time.

Preparedness means plan and rehearse, including stress testing key

infrastructure. However, the definition of infrastructure can be elusive. For example, work infrastructure seldom extends to employees' homes. Staff ordered to work at home become external network customers. Bandwidth to the home is seldom provided like bandwidth to workplaces, and some cannot upgrade. The terrain is changing. Preparedness is difficult, and it may be getting more challenging.

## RISK: TURNING WHITE SWANS BLACK

Preparedness requires learning about risk from

past experience. The United States has experienced multiple pandemics, including the relatively recent H1N1 and HIV/AIDS cases. Pandemic incidence might be increasing due to global travel, urbanization, human encroachment, environmental exploitation, and the emergence of new infectious diseases (an average of one each year over the past 30 years).<sup>9</sup> Epidemics and pandemics are always “on” in the sense that they are always coming. There may be notable differences, like a greater H1N1 impact among Asian countries or good protection therapeutics for HIV/AIDS. However, incidence was indicated. There would be epidemics and pandemics. Some countries such as Taiwan, Singapore, and Hong Kong have learned from SARS, which influenced fundamental preparedness changes. Others have been slower to learn.

COVID-19, per se, may not have been foreseen, but a pandemic was on. The COVID-19 pandemic started out

as a predictable “white swan” event, although it became a “black swan” with an extreme impact.<sup>10</sup> The public health community warned that a pandemic was inevitable, and the U.S. government has spent billions since 2005 to plan. Congressional hearings and Government Accountability Office reports have explored likelihoods and consequences.<sup>11</sup> Scientific literature projected pandemic infection, morbidity, and mortality, often with recommended steps to reduce risk. The COVID-19 pandemic became a black swan event the same way Hurricane Katrina did in 2005. There were failures at every level, and basic risk principles were violated.<sup>12</sup> The major causes were a lack of information, control, and time, as well as a failure to understand the precedence among them. A lack of information led to a lack of control, which led to a lack of time. Inadequate testing created ignorance of the pandemic’s extent and infection dynamics, which thwarted control of the COVID-19 spread. The only tool left was brute-force social distancing, which shut down most of the global economy.

Testing in the United States was slower than in South Korea, Singapore, and Taiwan, all of which did better.<sup>13</sup> The COVID-19 pandemic was under-resourced in the United States. Preparedness plans had addressed medical supply needs, but stockpiles were lacking. The priors did not include a major and fast-moving pandemic; it was assumed that no such pandemic would happen. SARS in 2003 and H1N1 in 2009 had modest impacts in most countries, creating complacency in the mistaken view that future pandemics would be mild.

Assumptions are the acceptance of risks. Risks show up at interfaces—interconnections with people, systems, or networks. The more interfaces there are, the more difficult it is to manage risk, and risk-management failures force reliance on powerful but blunt tools like social distancing that might mitigate pandemics but also stop work as interfaces are impaired or cut off. The result is surprise, often called

unintended consequences. Surprise is seen in how some U.S. grocers cannot restock their shelves. Interface failure is coupled with dependence on “lean inventory” that keeps minimal on-hand stock in local warehouses. Inventory is in the tightly coupled supply chain discussed later. Interfaces for communicating stocking information to suppliers are built around the assumption that suppliers operate under just-in-time delivery and only when needed. The COVID-19 pandemic crisis has made lean supply chains fragile.

Risk management requires engaging the risk ecology, which is an intersection of business, political, technological, and societal perils. Risk ecology lessons are predicated on interconnectedness as the key to modern systems.

### ANY WAR WILL SURPRISE YOU

A reporter once asked U.S. President Dwight Eisenhower about the likely outcome of a crisis. Eisenhower replied, “Any war will surprise you.”<sup>14</sup> In a 1906 speech at Stanford University, William James noted that people sometimes try to boost the salience of nonwar problems by making them the “moral equivalent of war.” Some have characterized the COVID-19 pandemic as war. This section explains why preparedness can be so difficult. To use another quote from Donald Rumsfeld, “You go to war with the Army you have, not the Army you might wish you have.”

First, even simple problems can surprise, as this true account suggests. A computer operations manager at a stock market data center in a New York skyscraper assured superiors of the facility’s state-of-the-art fail-over capability, portable generator for electricity, and so on. Then a superstorm hit, and the skyscraper’s basement, where the generator, telecom, and electric power circuits were located, filled with 13 ft of salt water. All circuits were fried, and the fault-tolerant computers failed. Everything had to be restored before the market reopened

in two days. Crews worked around the clock and got everything back up. Then a worker inadvertently crashed the system. The company missed the market reopening, lost customers, and damaged its brand. To avoid this kind of disaster in the future, the company spent millions of dollars on a roof-top generator and other backups. The lesson was to do this before the disaster. However, a decade later, water in the basement caused two of New York’s busiest medical centers to lose power in the wake of Hurricane Sandy.<sup>15</sup> The remediation plans were to prepare to fight the last war, not the next war. Plans did not address the availability of key people, as mentioned earlier.

In a crisis, surprise is revealed by everyday use. One example of this is a variant of the oft-discussed digital divide between haves and have-nots with Internet access that arose when its importance became clear. The COVID-19 pandemic illustrates a divide between “knows” and “know-nots.” In 2016, Cambridge Analytica showed the world that a Facebook app, *thisisyourdigitallife*, was gathering personal information from millions by exploiting weak security and privacy standards as well as layers of epistemic failures.<sup>17–20</sup> The disruptions caused by the COVID-19 pandemic could allow a foothold in commercial media platform news cycles and enable a version of Sarnoff’s law, which states that the political value of a message is proportional to the size of the audience and the frequency of the messaging. Crises produce large audiences that are ripe for modern advertising that came of age in the 1918 pandemic to boost morale for the World War I war effort.

Disruptive crises can facilitate controversial practices such as digital surveillance.<sup>25–27</sup> For-profit facial recognition service providers serve law enforcement, governments, and businesses with tools that have few checks on use or vetting for client acceptance.<sup>32</sup> Facial recognition services can match single images to databases of billions of images taken from social

media.<sup>28</sup> Crises like pandemics can fuel image capture while the public is frightened and suspends suspicions. BuzzFeed's assessment of 2,200 clients of facial recognition technology includes authoritarian regimes with questionable human rights records.<sup>34</sup> If COVID-19 draws attention from ethical and legal problems, those issues might become embedded in systems that arise from the pandemic.

COVID-19 might prove to be the opportunity to expand surveillance since organizations take advantage of the alienation and isolation of a population. Payback lies in capturing and monetizing personal information, either disclosed or otherwise. Crises can drive premature action while others are blamed. Some people are motivated by the perception of a threat more than by an actual threat. During crises there may be increased marketing of security systems. Because of the COVID-19 pandemic, ankle bracelet trackers and mobile fences are already in use in some countries. At the same time, there is little news about regulators who monitor behavior that might interfere with the status quo. Companies that promote these new technologies do not always address personal security and privacy issues effectively. Surveillance can backfire by alerting suspects to a location that is under surveillance and may be attacked (target prediction), so something that appears to be risk neutral can have consequences.

However, pandemics can expedite the sharing of data used to track exposure, implement quarantines, and conduct research. An earlier Ebola outbreak caused the U.S. government to relax some of the Health Insurance Portability and Accountability Act requirements to help with sharing, and similar actions have been taken during the COVID-19 pandemic response.<sup>16</sup> Crises can spur innovation by dealing with the unknowns. Ideas that would not have been tried might now be. This is a double-edged sword. Breakthroughs are possible, but foolish

mistakes due to inadequate due diligence can be costly and outweigh the benefits. Things can be more difficult than they look. For example, efforts to improve educational infrastructure could now be equivalent to the wet market where COVID-19 made the jump from animals to humans.

Zoombombing shows the pandemic crisis's impact on privacy and security.<sup>21,22</sup> Hard-learned lessons sometimes have to be relearned. There have been periodic pushes for technology in

Prior to COVID-19, technology-assisted education, especially online education, was limited to supportive roles despite experiments. In principle, online education is less expensive than the traditional model when the initial development costs can be amortized over a large enough student base (that is, they scale well). The lack of demonstrated educational improvement was offset by efficiency. When online is perceived to save money, it might be used. The COVID-19 pandemic

---

Breakthroughs are possible, but foolish mistakes due to inadequate due diligence can be costly and outweigh the benefits.

education, such as educational television and computer-assisted education. Halcyon claims such as personalized, self-paced, and self-directed instruction; immediate feedback; asynchronous delivery; and lack of bias have proved to be disappointing. They have not replaced traditional education, proving once again that technology seldom is a panacea.

The pandemic has caused education to embrace online classes because there was no other choice. That being said, the net results thus far are not clear. Students and teachers who like online collaboration make it work. The accomplishments of online learning in the COVID-19 pandemic era would not have been possible even 10 years ago. The online comfort level of the students could be a factor, but it is not well understood. The effects of differences in comfort levels among or between students and teachers are theoretically important but unknown. Some like online learning while others do not. Risk assessment is difficult because it is unclear what will or can happen. The risk ecology itself is destabilized, allowing both innovation and mistakes.

This dynamic produces a conundrum for the risk ecology in education.

triggered the widespread use of online education because there was no alternative. The verdict awaits. However, there have been a few offers of tuition remission to reflect whatever cost savings institutions realized or the degradation in quality some experienced. Some tuition payers feel they have been charged full price for impoverished service. The conundrum is whether to admit the problems by refunding tuition or forge ahead with online education as mainstream, the "new normal." Aside from the risks involved in contests over tuition remission, important risk issues include reputation and value adds.

The consequences of new business models are sometimes scrutinized less than they should be, especially regarding quality of service and satisfaction of expectations. Serious potential problems can be ignored due to a lack of awareness and understanding. In the case of the push to move to online education, older strategies like interactive, duplex environments (teleconferencing) or rectified or simplex systems (podcasting) might be ignored in favor of more sophisticated services. Such approaches may have been debugged through decades of use, entail minimal expense, and



carry little or no privacy and security risks for participants. However, such suggestions can be met with the response that these capabilities are embedded in commercial platforms. Survival mode can make choices narrow and comparisons difficult. Nevertheless, the allure of sophisticated alternatives is strong, especially when coupled with the loss leader of free service. In fact, free can turn users into exploitable products by exposing them to risk. They have their place as part of a socially responsible, measured, and informed risk ecology. They are not a “go-to” solution.

Dependence is often tied to progress. It is said that necessity is the mother of invention, but invention is also the mother of necessity. People come to depend on inventions. Dependence sneaks up as new circumstances produce cumulative changes over time. Most of the time, things run normally. The desire to make them robust with fault-tolerant design, backups, redundancy, and so on makes them brittle. When disruptions occur, flexibility that maintains the essential is needed. It is not possible to build integrated systems that are simultaneously flexible and robust. During crises, systems must change from robust to flexible, often quickly. With enough complexity, greater flexibility requires controlled disintegration. The automation paradox in cockpits occurs when pilots become dependent on automation for safe aircraft operation, although automation failure requires pilots to know what to do. Sometimes they do not know.

Another example is problems with state unemployment systems that have broken down during the COVID-19 pandemic. One state had massive difficulty in handling unemployment claims, even though it had modernized its system after the Great Recession of 2008–2010. The modernized system was designed to handle only the level of unemployment experienced previously, since the new system requirements assumed unemployment would never be greater than the

Great Recession. If the requirements were tied to the Great Depression of the early 1930s, the system would have handled the load. This is another form of preparing to fight the last war.

Finally, there is a link between IT and tightly coupled systems. This article cannot provide a full explanation of these coupled systems, but the attention paid to this topic is likely to increase. Coupling is part of the risk ecology and drives risk. The Allied victory in World War II (WWII) often points to technology like nuclear weapons, code-breaking machines, radar, the Douglas DC-3 planes, the Jeep, antibiotics, and so on. Equally important but discussed less often is how broken-down machines (trucks, tanks, ships, airplanes, and weapons) could be fixed quickly, which is an advantage in mechanized war. These technologies were part of loosely coupled systems and amenable to repair skills learned on farms and in factories served by erratic and often slow supply chains.

After WWII, loosely coupled systems gradually gave way to tighter coupling as knowledge and new technology, especially IT that helped integrate disparate information sources, became more important. Integration produced highly capable systems. A good example is an engine management systems (EMSs) for vehicles with internal-combustion engines. EMSs improved engine performance, fuel economy, and longevity. Without them, meeting progressively stricter emissions standards would not have been possible. However, these tightly coupled systems require specialized know-how, expensive software-controlled diagnostic tools, specialized repair tools, and an elaborate supply chain to be fixed. Few people can repair these engines.

Systems increasingly depend on tightly coupled software. Brooks noted that, with the IBM 360 operating system, fixing tightly coupled systems can introduce new errors.<sup>23</sup> Tightly coupled systems are susceptible to disruptions and difficult to test. During the Cold

War, people feared that a Soviet electromagnetic pulse from high-altitude nuclear blasts would cripple tightly coupled U.S. weapon systems while the old-fashioned, loosely coupled Soviet systems remained unharmed. Tightly coupled weapons systems can be difficult or impossible to test without an actual war. Yet tightly coupled systems of all kinds have become ubiquitous and essential. The utility infrastructure increasingly depends on system control and data acquisition (SCADA) networks, often using the Internet. Crucial services, including banking, air travel reservations, unemployment benefits, and other government systems, now operate on systems that are tightly coupled to software developed decades ago.<sup>24</sup> This issue was behind much of the Y2K Problem.

Sociologist Charles Perrow said that tightly coupled systems are prone to “normal accidents.”<sup>29</sup> These are not aberrations, and they are inevitable. Tightly coupled supply chains for toilet paper and food have been made famous by the COVID-19 pandemic. Shortages in one supply chain can be offset by surpluses in others. These supply chains were not designed; they evolved. They have never been redirected or stopped, and now they have must be redirected while, in some cases, the pandemic has stopped them. There has been a push to redirect the commercial toilet paper and food supply chains toward residential use. For a time, there were no cargo ships from China in the ports of San Pedro (Los Angeles and Long Beach, California), which are primary Chinese entrepôts. As inventory moved from stocks in warehouses to flows aboard “lift,” inventory in transit became the only form of inventory. As discussed in lean grocery inventory earlier, tightly coupled supply chains cannot be redirected easily. It is not clear what it takes to restart them if they stop. Many unexpected deadly embraces and other problems are likely.

Computerized, tightly coupled systems have become vital to society and the economy. It is increasingly difficult

to predict what failures in such systems might mean. Poor preparedness decisions can carry great risk, with consequences that outlast the crises that prompted them. As noted before regarding privacy and security concerns, social costs can be high. Crises can accelerate already hasty decision processes. When a crisis momentarily provokes attention, the quality of decisions can decline.

## RISK FATIGUE

Risk fatigue causes people to turn an ineffective (but not entirely blind) eye to crises, ignoring likelihood or even certainty, dismissing the risk ecology, and preparing to fight the last war. Risk fatigue is normal. While claims that an emergency disruption could not have been foreseen are preposterous in one sense (it's on), they are legitimate in that it is impossible to see the future. An analogous event in recorded history, especially in living memory, proves that such things can happen. Sometimes it is possible to know how frequently they occur and how disruptive they can be. This section discusses the causes of risk fatigue, showing that it is to be expected.

Over time, especially for rare events, the vigilance of individuals and organizations atrophies without recurring triggers (for example, close calls) to heighten awareness.<sup>30</sup> If nothing untoward happens even with triggers, it can take effort and energy to avoid complacency. Constant admonitions by authoritative individuals and organizations to prepare for a crisis wear off as people become comfortable with the notion of the crisis coming. In the case of pandemics, calls for preparation can have the opposite effect by appearing to be overwrought.

Infrequency can cause risk fatigue. Infrequent events, irrespective of intensity, are forgotten. Attention may be paid to emergency preparedness during and immediately after infrequent events, but the people involved disappear and memory fades. Preparation becomes more "real" than the

crisis being prepared for. People forget without reinforcement. Strangely, high frequency can also cause risk fatigue as events become routine.

There is also social amplification of risk, in which some risks receive more attention than is called for.<sup>31</sup> The world seems split between infrequent but disruptive events (great tornados, hurricanes, thunderstorms, earthquakes, and so forth, especially in regions where such events seldom occur) and routine events that are handled regularly. Big but infrequent events "never happen" while frequent events are not worth discussing. Any place with frequent huge events is uninhabitable.

Staff designated to be prepared are often involved with IT because, in most organizations, they have experience with systems. Others turn to them to lead preparedness and response. The wireline telephone system was hardened to keep working when other utilities failed. It had its own electrical capability and no end-user data storage. Similarly, large computer systems had uninterruptible power supplies and backup for data storage, while end users had limited local data. As distributed technologies proliferated (cellular telephony, personal computers, and so on), dependency grew as power and data storage exposure increased. IT has become more central, from utility management (namely, SCADA networks) to transaction processing and storage of organizational and personal data. The assurance of robustness falls on IT managers since IT departments have functioned as risk management pioneers in most organizations. The IT department is presumed to have risk expertise.

A particularly problematic task for IT functions is to get others to understand the dynamics of system integration. As noted, risks occur at interfaces that proliferate under system integration. The dream of an integrated system as organizational panacea is old. The Urban Information Systems Inter-Agency Committee (USAC) program ended in 1977, after

spending more than US\$26 million (more than US\$170 million today) to build integrated municipal information systems.<sup>32</sup> Nearly 80 teams of municipalities, computer companies, and universities submitted proposals, and six cities were selected to build integrated systems or subsystems. Ten federal agencies led and paid for USAC. The lead agency was Civil Defense (CD), which was frustrated because cities did not replenish perishable supplies in the Cold War emergency shelters for which CD was responsible. Since integrated, computerized information systems were proposed as the solution to this problem, shelter maintenance became part of routine operational information updates. USAC advanced municipal information systems, but CD's dream was not met. It turned out that integrated systems were not worth the trouble. CD itself eventually fell apart before the end of the Cold War.<sup>33</sup>

## TIPS

From the preceding, we provide the following three tips for preparedness.

1. *Pick your battles:* Managing risk is about the future of present decisions. Use "failure imagination" to determine the worst outcome. Manage expectations before and during the crisis. Beforehand, get people to understand that the goal is not to have business as usual during the crisis. Rather, it means prioritization: deciding in advance what will be attended to and what will be ignored. Nobody does more with less. They do less with less. The primary job is to decide what subset of the current will be done and how to transition to that. Policy made during a crisis is temporary. After the crisis, everything returns to the status quo ante bellum. Although lessons learned during the crisis might influence future decisions, there is no replacement for due diligence. A plan to align

authority and responsibility during the crisis is especially important if it is different than normal.

2. *Plan for controlled disintegration:* Systems (technology, supply chains, governance) are becoming more complex and tightly coupled, failures cascade, and failure costs escalate. System integration can make things worse for emergency preparedness. Consider controlled disintegration. Imagine the new, not just the things that everybody knows. For example, during the COVID-19 pandemic, work was often relocated to homes. This might be common in the future or extend to other areas, such as telemedicine. Test for the full capacity needed for such moves. Turn things off to highlight coupling and dependencies. Decide priorities for systems and subsystems in crisis and how they will be maintained. Low-priority items must be decoupled from integrated systems so their problems do not affect essential functions. This must be planned and practiced. Risk is in interfaces. In crises, it must be reduced by becoming less tightly coupled. A firewall must be built against cascading failures by simplifying around crisis essentials. What is important during a crisis may be different from normal. Embrace triage. Recognize that falling over is falling back. Plan how key suppliers, vendors, consultants, utilities, and customers will function in crisis and how to handle those situations.
3. *Test your assumptions:* Assumptions are where 99% of failures start. They are risks taken. Repeatedly and vigorously test assumptions to uncover unknowns and overcome risk fatigue. The risk ecology is

always changing as well as assumptions and opportunities to mitigate change. Past risks can fade away while new ones come. The COVID-19 pandemic reveals assumptions that were not tested thoroughly. Soon there will be a surplus of ventilators, yet last year everyone assumed it would take years to produce so many of them.

**M**ultiple cross-cutting threads are presented in this article. Crises are not homogeneous. Some crises impact facilities (for example, fires), others impact people (such as pandemics), while others impact both (in particular, earthquakes). Crises are on—they will happen. There is some evidence that they are becoming more frequent and more complex. There rarely are enough preparedness resources. The key to preparing to function in a crisis with complex systems is to simplify at crisis time, reducing obstacles like privacy and security, knowing that these will be reengaged when the crisis passes. As urgency rises, so does expediency. ■

## REFERENCES

1. J. G. March and H. A. Simon, *Organizations*. New York: Wiley, 1958.
2. D. Rumsfeld, *Known and Unknown: A Memoir*. New York: Penguin, 2010.
3. "Sigma 1/2020: Data driven insurance," Swiss Re Institute, Sigma, Zürich, Switzerland, Feb, 2020, p. 28. [Online]. Available: <https://www.swissre.com/institute/research/sigma-research/sigma-2020-01.html>
4. L. Lightbody and M. Fuchs, "Every \$1 invested in disaster mitigation saves \$6: Spending to reduce risk saves lives and creates jobs, key study finds," Pew Charitable Trust, Philadelphia, PA, Jan. 11, 2018. [Online]. Available: [https://www.pewtrusts.org/en/research-and-analysis/articles/2018/01/11/every-\\$1-invested-in-disaster-mitigation-saves-\\$6](https://www.pewtrusts.org/en/research-and-analysis/articles/2018/01/11/every-$1-invested-in-disaster-mitigation-saves-$6)
5. "Natural hazards, unnatural disasters: The economics of effective prevention," World Bank and UN, Washington, D.C., 2010, pp. 10, 14. [Online]. Available: <http://documents.worldbank.org/curated/en/620631468181478543/pdf/578600PUB0epi2101public10BOX353782B.pdf>
6. "ND-Gain Country Index," Notre Dame University Global Adaptation Initiative (ND-GAIN), South Bend, IN, 2017. Accessed on: May 15, 2020. [Online]. Available: <https://gain.nd.edu/our-work/country-index/>
7. Wikipedia, "List of countries by natural disaster risk," 2017. Accessed on: May 15, 2020. [Online]. Available: [https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_natural\\_disaster\\_risk](https://en.wikipedia.org/wiki/List_of_countries_by_natural_disaster_risk)
8. "World First Aid Day: Red Cross Red Crescent calls for expansion of life-saving skills for everyone, everywhere," IFRC, Geneva, Switzerland, Sept. 12, 2014. [Online]. Available: <https://www.ifrc.org/en/news-and-media/press-releases/general/world-first-aid-day-red-cross-red-crescent-calls-for-expansion-of-life-saving-skills-for-everyone-everywhere/>
9. Institute of Medicine, Forum on Microbial Threats, *Microbial Evolution and Co-Adaptation: A Tribute to the Life and Scientific Legacies of Joshua Lederberg: Workshop Summary (Infectious Disease Emergence: Past, Present, and Future)*. Washington, D.C.: National Academies Press, 2009, p. 5. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/20945572>
10. H. R. Morgan, "Why the coronavirus may be a black swan event," Inc., Feb. 29, 2020. [Online]. Available: <https://www.inc.com/heather-r-morgan/why-coronavirus-is-a-black-swan-event-we-might-actually-need.html>
11. "Lessons from the H1N1 pandemic should be incorporated into future planning," Government Accountability Office, Washington, D.C., GAO-11-632, June 27, 2011. [Online].

- Available: <https://www.gao.gov/new.items/d11632.pdf>
12. R. N. Charette, *Software Engineering Risk Analysis and Management*. New York: McGraw-Hill, 1989. [Online]. Available: <https://www.amazon.com/Software-Engineering-Analysis-Management-ENGINEERING/dp/0070106614>
  13. J. Asquith, "Encouraging Outlook—In Taiwan, Singapore and South Korea life is continuing without lockdowns," *Forbes*, Apr. 1, 2020. [Online]. Available: <https://www.forbes.com/sites/jamesasquith/2020/04/01/positive-outlook-in-taiwan-singapore-and-south-korea-life-is-continuing-with-relative-normality/#5b0440eb7335>
  14. C. Allen, *Eisenhower and the Mass Media: Peace, Prosperity, and Prime-Time TV*. Chapel Hill, NC: Univ. of North Carolina Press, 1993, p. 91.
  15. "What caused generators to fail at NYC hospitals?" CBS News, Nov. 2, 2012. [Online]. Available: <https://www.cbsnews.com/news/what-caused-generators-to-fail-at-nyc-hospitals/>
  16. J. Davis, "HHS issues limited waiver of HIPAA sanctions due to coronavirus," Health IT Security. [Online]. Available: <https://healthitsecurity.com/news/hhs-issues-limited-waiver-of-hipaa-sanctions-due-to-coronavirus>
  17. C. Wylie, *Mindf\*ck: Cambridge Analytica and the Plot to Break America*, New York: Random House, 2019.
  18. R. McNamee, *Zucked: Waking up to the Facebook Catastrophe*. Baltimore, MD: Penguin, 2019.
  19. K. H. Jamieson, *Cyber-War: How Russian Hackers and Trolls Helped Elect a President*. London: Oxford Univ. Press, 2018.
  20. H. Berghel, "New perspectives on (Anti)Social Media," *Computer*, vol. 53, no. 3, pp. 77–82, Mar., 2020. doi: 10.1109/MC.2019.2958448.
  21. A. Zimmerman and C. Veiga, "As NYC bans Zoom for online learning, some schools pause live instruction," Chalkbeat, Apr. 6, 2020. [Online]. Available: <https://chalkbeat.org/posts/ny/2020/04/06/nyc-schools-zoom-ban/>
  22. N. Anderson, "'Zoombombing' disrupts online classes at University of Southern California," *Washington Post*, Mar. 25, 2020. [Online]. Available: <https://www.washingtonpost.com/education/2020/03/25/zoombombing-disrupts-online-classes-university-southern-california/>
  23. F. Brooks, *The Mythical Man-Month*. Reading, MA: Addison-Wesley, 1975.
  24. A. Lee, "Wanted urgently: People who know a half century-old computer language so states can process unemployment claims," *CNN*, Apr. 8, 2020. [Online]. Available: <https://www.cnn.com/2020/04/08/business/coronavirus-cobol-programmers-new-jersey-trnd/index.html>
  25. D. Carroll, "Internet worm linked to San Francisco man," *The Harvard Crimson*, Feb. 25, 2009. [Online]. Available: <https://www.thecrimson.com/article/2009/2/25/internet-worm-linked-to-san-francisco/>
  26. O. Thomas, "The person behind a privacy nightmare has a familiar face," *San Francisco Chronicle*, Jan. 22, 2020. [Online]. Available: <https://www.sfchronicle.com/business/article/The-person-behind-a-privacy-nightmare-has-a-14993625.php>
  27. L. O'Brien, "The far-right helped create the world's most powerful facial recognition technology," *Huffington Post*, Apr. 7, 2020. [Online]. Available: <https://www.huffpost.com/entry/>
  28. R. Mac, C. Haskins, and L. McDonald, "Clearview's facial recognition app has been used by the Justice Department ICE, Macy's, Walmart, and the NBA," *BuzzFeed*, Feb. 27, 2020. [Online]. Available: <https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement>
  29. C. Perrow, "Normal accident at three Mile Island," *Society*, vol. 18, no. 5, pp. 17–26, 1981. doi: 10.1007/BF02701322.
  30. W. R. Freudenburg, "Nothing recedes like success? Risk analysis and the organizational amplification of risks," *Risk, Issues Health Safety*, vol. 3, no. 1, pp. 1–135, 1992.
  31. R. E. Kasper et al., "Social amplification risk: A conceptual framework," *Risk Anal.*, vol. 8, no. 2, pp. 177–187. doi: 10.1111/j.1539-6924.1988.tb01168.x.
  32. K. L. Kraemer and J. L. King, "A requiem for USAC," *Policy Anal.*, vol. 5, no. 3, pp. 313–349, 1979.
  33. D. Garrison, *Bracing for Armageddon: Why Civil Defense Never Worked*. Oxford, U.K.: Oxford Univ. Press, 1976.
  34. C. Haskins, R. Mac, and L. McDonald, "Clearview AI wants to sell its facial recognition software to authoritarian regimes around the world," *BuzzFeed*, Feb. 5, 2020. Accessed on: May 15, 2020. [Online]. Available: <https://www.buzzfeednews.com/article/carolinehaskins1/clearview-ai-facial-recognition-authoritarian-regimes-22?bfsource=relatedmanual>

**HAL BERGHEL** is a Fellow of the IEEE and ACM and a professor of computer science at the University of Nevada, Las Vegas. Contact him at [h1b@computer.org](mailto:h1b@computer.org).

**ROBERT N. CHARETTE** is the founder of ITABHI Corporation. Contact him at [ncharette@ieee.org](mailto:ncharette@ieee.org).

**EDWARD G. HAPP** is an executive fellow at the School of Information, University of Michigan. Contact him at [ehapp@umich.edu](mailto:ehapp@umich.edu).

**JOHN LESLIE KING** is a W.W. Bishop collegiate professor at the School of Information, University of Michigan. Contact him at [jlking@umich.edu](mailto:jlking@umich.edu).





# 21 Years of Distributed Denial-of-Service: Current State of Affairs

**Eric Osterweil and Angelos Stavrou**, George Mason University

**Lixia Zhang**, University of California, Los Angeles

*The Internet's features and capacity have evolved, but is the nature of its security noticeably better? We examine the fundamental nature of distributed denial-of-service (DDoS) attacks and the state of the union of our defenses in today's DDoS wars.*

In 1999 (21 years ago), malware called Trin00<sup>1</sup> compromised a set of computers and then took down a network at the University of Minnesota. This event marked the birth of volumetric distributed denial-of-service (DDoS) attacks from robot networks (botnets). While earlier attacks exist in anecdotes and recollections, this documented case sets a lower bound on the date of birth: 21 years. The

features and capacity of the Internet have evolved a lot since then, but is its security disposition demonstrably better? Trin00 used hundreds (possibly thousands) of compromised machines (bots), but today conventional botnet sizes have been seen in the millions. In relative terms, Trin00 may not seem like such a large botnet. However, this underscores that historical attack sizes are relative, and raw numbers alone do not tell the tale. Moore's law and bandwidth increases makes comparing attack volumes (bits per second) from the

past to today (or tomorrow) apples-to-oranges comparisons. Consider that gigabit attacks in 2000 were considered staggering, but only because they rivaled the capacity of the infrastructure of the time. An unfortunate state of affairs is that it has always been easier to gain attack capacity than defensive capacity.<sup>2</sup> DDoS is an asymmetric threat with an impedance mismatch between attackers and defenders.

The gap between adversaries' barriers to attack and the price to defend has always been large, but it is growing, and the status quo does not paint a pretty picture for the future



of Internet service security. In this article series, we want to sound an alarm and issue a call to action; we must discover the fundamental enablers of DDoS, and we must use these to craft efficient defenses. We feel it is time to reexamine the principles that underlie the problem space. In this two-part article, we begin by examining the fundamental nature of DDoS: reasons why our networks are susceptible, the anatomy and nature of today's DDoS attacks, and the state of the union of our defenses in today's DDoS wars. In our article's second part (in the August issue of *Computer*), we explore remediations and the evolution needed to systematically enhance the Internet and address the principles that enable DDoS.

## WE ARE VICTIMS OF OUR OWN SUCCESS

The Internet has blossomed with complex and diverse network applications and services that bind our social lives, implement complex tasks, and facilitate communications, all while streamlining end users' experiences. Critical to this success has been protocol layering and abstraction (where protocols encapsulate and obscure their state from each other). Network applications sit above the transport layer, which sits above the network layer. Indeed, layering has been a central tenet of the Internet's evolvable architectures. However, it also has hidden vulnerabilities that many DDoS attacks now capitalize on. As defenders against DDoS attacks, our fundamental challenge is the onus of scrubbing attack traffic away from legitimate traffic, using deep packet inspection (DPI). The distinction of which traffic is part of an attack is often only visible at the application layer (above both the network and transport layers). For example, what network-layer information should management tools use to determine which Domain Name System (DNS) queries are real and which

are participating in a reflection attack? Which Network Time Protocol (NTP) command is legitimate and which is part of an attack? Must it fall to the application programmer and operators to build custom logic to differentiate whether a memcached<sup>3</sup> query is from a genuine application or part of an attack? Or which

starve servers of resources and can be hard to detect. Some are volumetric, which send overwhelming volumes of traffic that congest network links and overload servers and are hard to stop even though they are detectable by nature. In this article, we focus our discussions on these two types.

---

Our fundamental challenge is the onus of scrubbing attack traffic away from legitimate traffic.

HTTP client is trying to keep a needed connection alive and which is starving the server for resources?

Compounding the opacity across layers, network traffic is now often encrypted. A recent operational report of large-scale measurements stated that the Secure Sockets Layer (SSL) "is [sic] majority of traffic in [North America] by February 2019."<sup>4</sup> The necessary computational complexity, volume of traffic, and growing use of encryption often render common operational network tools ineffective in defending against attacks. When application payloads are embedded (that is, encrypted), they require multiple layers of computationally expensive decoding while exposing sensitive material. For example, performing DPI on an HTTPS flow requires decryption of the flow. Further, that also requires escrow of the end site's Transport Layer Security (TLS) private key (to terminate and inspect the embedded flow). Internet protocol layering and encryption have severely complicated scrubbing at the network layer. In short, this is high cost and low return, and it is time to investigate the fundamentals of this problem space.

## A GLIMPSE AT THE ANATOMY OF DDOS

DDoS comes in a variety of flavors. Some are called "low-and-slow," which

### Volumetric DDoS today

Today, DDoS threats are asymmetric: it is virtually free for attackers to acquire massive network capacities for their DDoS attacks, and they frequently use multiple techniques, tactics, and procedures (TTPs) at the same time. By contrast, detecting and mitigating DDoS attack traffic (for example, "packet love") requires investment in expensive infrastructure and network bandwidth (capacity). The largest recorded DDoS attacks have used source address spoofing (such as sending packets with deliberately falsified addresses) as part of their TTPs. One form of spoofing attack amplifies its volume by bouncing (or "reflecting") small queries off of Internet services to elicit larger ("amplified") responses, which are then reflected to spoofed addresses (in other words, victims). These are called *reflective amplification attacks* or just *reflector attacks*. For example, in 2016, the first publicity around a terabit attack came from an assault on a hosting provider called OVH,<sup>5</sup> and it reached this volume by using spoofed addresses in a reflector attack. A larger attack on Dyn<sup>6</sup> surpassed this volume, again in 2016, using source address spoofing. In short, the largest DDoS attacks seen today depend on address spoofing as part of their TTPs, though they have not always leveraged an amplification factor.

The increasingly relative ease of acquiring large volume attack sources has

elevated the appeal of volumetric DDoS attacks to adversaries. Traffic may be DNS queries, Simple Network Management Protocol queries, NTP queries, memcached queries,<sup>3</sup> or others. Some attacks also use spoofed Transmission Control Protocol (TCP) control traffic carrying large data payload or use the TCP session itself as an amplification vector by orchestrating torrents of reset packets or data payloads via the TCP PSH option.

### Server-side resource exhaustion attacks

While many headlines and defenses focus on the size of DDoS attacks, there continue to be many attacks above the transport layer. Common examples of

perhaps more troubling is the fact that their detection and remediation more clearly requires additional state information above the network layer.

### MITIGATION: STATE OF THE UNION, TODAY

Mitigation providers often do distribute their services, often called *scrubbing centers*, around the world and across the Internet's topology. However, with attack sources sometimes numbering in the millions, scrubbing centers each inevitably need to mitigate attack traffic from growing numbers of well-provisioned botnets. Scrubbing uses DPI, thereby adding computation overhead to the network/transit overhead. This frames the fundamental impedance

applications and use of encryption. Our mitigation techniques are predicated on matching mitigation bandwidth to ever-growing aggregate distributed attack volumes, and we need a different/more distributed solution. For example, in 2015, the Defense Advanced Research Projects Agency announced a call for "Extreme DDoS Defense" that included a solicitation to "[disperse] cyber assets (physically and/or logically)."<sup>10</sup>

Some techniques to disperse network-based remediation focus on using network-layer routing, like Border Gateway Protocol's (BGP's) FlowSpec, remotely triggered black-holing, and others. However, without the necessary application-level expressiveness, this can unfortunately lead to collateral damage to well-behaving (nonattack) sources that happen to be on the same network (such as in the same BGP prefix) as attackers.

Another network-layer defense, called Internet Protocol (IP) anycast, uses BGP routing to replicate services. Anycast allows operators to position services near clients and provides redundancy. However, Internet Architecture Board RFC 7094<sup>11</sup> describes some known limitations: "IP control packets from a DNS client may initially be routed to one anycast instance, but subsequent IP packets may be delivered to a different anycast instance." Recent work<sup>12</sup> examined the DNS anycast root server system while under sustained DDoS and concluded that there is a "need to understand anycast design for critical infrastructure, paving the way for future study in alternative policies that may improve resilience."

### Volumetric state of the union

The volumetric state of the union—volumes of attack traffic versus carriers' and providers' provisioned capacities—paints a similarly disconcerting picture. Service providers (SPs) buy transit in gigabits per second (Gbit/s) links in multiple locations from multiple carriers. Internet exchange points and carrier capacity are also often offered in Gbit/s. Large carriers' global

The increasingly relative ease of acquiring large volume attack sources has elevated the appeal of volumetric DDoS attacks to adversaries.

such attacks leverage protocol aspects in the SSL, TLS, or even at the HTTP/S layer. These nonvolumetric attacks can also be crippling without a DDoS defense system and can bring Internet services down with far fewer resources.

Perhaps the earliest known resource exhaustion attacks were those that abused the TCP itself, SYN flood attacks. These DoS attacks have been used in the wild since at least 1996,<sup>7</sup> though they were not always distributed. Attacks like these were initially intended to exhaust servers' resources and were neither volumetric nor stealthy (low and slow).

One of the early examples of low-and-slow attacks was Slowloris,<sup>8</sup> where a relatively small number of stateful HTTP queries would hold connections open on web servers and thereby exhaust their ability to answer other (legitimate) clients. Other exhaustion attacks exploit TLS's cryptographic key negotiation.<sup>9</sup> In these types of attacks, the raw numbers of attacking clients and traffic are not as spectacular as volumetric DDoS, but

mismatch: distributed attacks versus relatively centralized mitigations. As just an illustration, we present three examples that the state of the Internet can be categorized and evaluated: architectural, volumetric, and economic.

### Fundamentals of the state of the union

Our reliance on DPI for detection and remediation has resulted in increasing dependence on keeping our defenses in large computation/network capacity data centers. For example, with reflector attacks leveraging application-level semantics (for example, NTP's monlist and memcached's GET) and the increased use of TLS, terminating and interpreting traffic has necessitated backhauling traffic to DPI in scrubbing centers. This has framed a fundamental asymmetry: large volumes of attack traffic from more sources with increasingly better provisioned networks versus fewer and centralized remediation. This asymmetry is further exacerbated by the increased complexity of web

aggregate capacity may approach, and in some cases achieve, terabits per second (Tbit/s). However, this does not mean any given ingress point to a carrier's network is itself a Tbit/s link. Generally, aggregate capacity in Tbit/s is a summation of router/regional capacities (Gbit/s). However, the aggregate attack traffic of the largest DDoS attacks is already over 1 Tbit/s. In an aggregate view, a recent observation from operational measurements quotes that "attacks [are] growing in size faster than network growth."<sup>3</sup>

Unfortunately, often the aggregate capacity is not near attack sources, and it can be topologically very far from attack sources. While the volume of observed DDoS attacks has already crippled critical infrastructure, the potential sizes of attacks is far worse than anything that we have seen to date. "The Internet's capacity attenuates the total throw weight a DDoS attack can generate; the farther a target is from components of a network, the less traffic that will make it across any congested links between the target and the attack source."<sup>13</sup> In other words, this can result in service degradation and outages to other Internet services whose traffic shares congested routing infrastructure as they become collateral damage. This was also noted during the Spamhaus/Cloudflare DDoS of 2013.<sup>14</sup> When attack sources are topologically far from mitigation, their traffic is backhauled across transit and peering infrastructure to scrubbing centers causing terabits of attack traffic to potentially be routed to gigabit scrubbing centers. Even in the case of high-capacity scrubbing centers, the centralized nature of the mitigation enables attack traffic to permeate network links far from the sources of attack.

The largest DDoS attacks that we have seen are already larger than the provisioned capacity of many of the large providers' and carriers' capacities. In 2016, the U.S. Department of Homeland Security started a program called DDoS Defense, whose starting position was that "one day" DDoS could swell to 1 Tbit/s.<sup>15</sup> By 2017, the largest

DDoS attacks had already reached that, and in 2018 DDoS attacks quantifiably exceeded that, as shown in Figure 1.<sup>16</sup> Recent work has estimated that the Internet-wide capacity to launch volumetric reflective amplification DDoS attacks is "two orders of magnitude larger than the Dyn attack."<sup>17</sup>

### Economic state of the union

Using SPs' outlays to protect against DDoS also paints a grim picture. In

is growing. Furthermore, there has also been a DDoS-for-hire (sometimes known as a *booter*) grey-market for roughly a decade.

The motivations for launching DDoS attacks can be diverse. For example, in 2015 the hacktivism group Anonymous threatened to—and then did—launch a DDoS attack against the DNS root server system.<sup>16</sup> The stated goal of this attack was to disrupt all transactions on the Internet by rendering the DNS

## The motivations for launching DDoS attacks can be diverse.

2000, DDoS attacks on Yahoo, eBay, and several other major Internet services led the news and raised alarms. Now, almost 21 years later, protection rackets exist in gaming spheres. Online gaming and gambling sites are frequently held hostage for ransom by DDoS threats,<sup>18</sup> and sometimes attacks are launched simply to gain gaming advantages. Generally, all online services today need DDoS protection, and companies expect to pay for defensive protections against inevitable DDoS attacks. The DDoS mitigation market was US\$1.94 billion in 2018 and

inoperable. While unsuccessful, this attack illustrates that sometimes DDoS attacks are launched to wreak Internet havoc and do not have a specific target.

### ADDRESSING ROOT CAUSES

Internet providers and clients seek protection from DDoS attacks in advance of, during, and subsequent to them. However, there are no official authorities to enforce or remedy DDoS. There is no government mandate or Internet regulatory body that has the authority or is even in a position to offer remediation

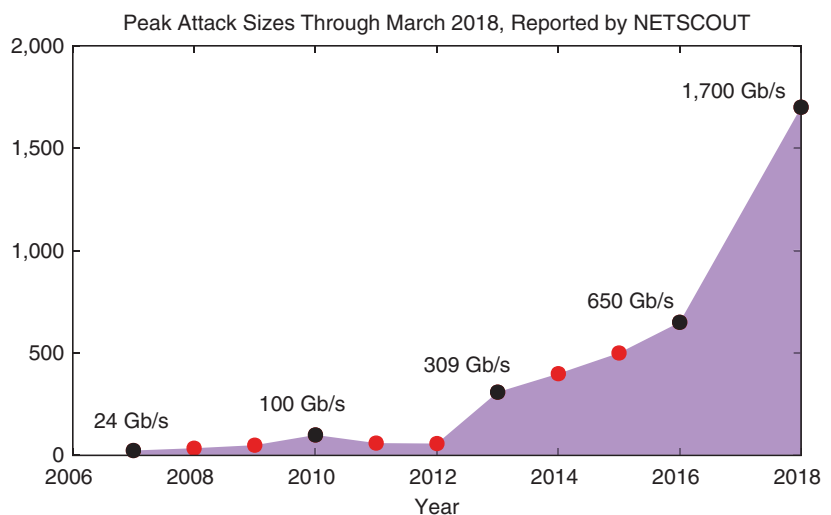


FIGURE 1. The peak attack sizes through March 2018.<sup>16</sup>



or help, and not everyone would want to bestow such global authority to an organization. As of today, we must pay for help, and we have privatized our defenses. This begs the unfortunate question: what would a proper remedy to DDoS attacks even look like? Considering that DDoS traffic often originates from multiple countries, may transit through separate jurisdictions, may lie about where it comes from (spoofing), and the fact that the Internet's infrastructure is operated by private corporations, is it even feasible that any entity could provide official remedies to DDoS attacks? We feel the biggest challenge, which must be addressed first, is to examine the root-cause vulnerabilities that enable DDoS attacks and then to develop mitigation techniques.

In next month's column, we will detail the approaches that are used to combat DDoS today, examining scenarios and conditions under which they perform well. We will also explore opportunities and directions for basic and applied research to address the fundamental attack vectors that DDoS attacks exploit. ■

## REFERENCES

1. "CERT incident note IN-99-04," CERT Coordination Center, Pittsburgh, PA, 1999. [Online]. Available: [https://web.archive.org/web/20081115163511/http://www.cert.org/incident\\_notes/IN-99-04.html](https://web.archive.org/web/20081115163511/http://www.cert.org/incident_notes/IN-99-04.html)
2. G. Huston, "Why is securing the Internet so hard?" in *Proc. Asia Pacific Regional Internet Conf. Operational Technologies*, 2019.
3. "Memcached DDoS explained," Akamai, Cambridge, MA. [Online]. Available: <https://www.akamai.com/us/en/resources/our-thinking/threat-advisories/ddos-reflection-attack-memcached-udp.jsp>
4. C. Labovitz, "Internet traffic 2009–2019," in *Proc. Asia Pacific Regional Internet Conf. Operational Technologies*, 2019.
5. E. Kovacs, "Hosting provider OVH hit by 1 Tbps DDoS attack," *SecurityWeek*, 2016. [Online]. Available: <https://www.securityweek.com/hosting-provider-ovh-hit-1-tbps-ddos-attack>
6. S. Mansfield-Devine, "DDoS goes mainstream: How headline-grabbing attacks could make this threat an organisation's biggest nightmare," *Netw. Secur.*, vol. 2016, no. 11, pp. 7–13, 2016. doi: 10.1016/S1353-4858(16)30104-0.
7. "CERT advisory CA-1996-21 TCP SYN flooding and IP spoofing attacks," CERT Coordination Center, Pittsburgh, PA, 1996.
8. R. Hansen, J. Kinsella, and H. Gonzalez, "Slowloris HTTP DoS," *h.ckers*, 2009. [Online]. Available: <https://web.archive.org/web/20150426090206/http://h.ckers.org/slowloris>
9. C. Castelluccia, E. Mykletun, and G. Tsudik, "Improving secure server performance by rebalancing SSL/TLS handshakes," in *Proc. ACM Symp. Information, Computer and Communications Security*, 2006, pp. 26–34.
10. J.M. Smith, "Extreme DDoS defense (XD3)," Defense Advanced Research Projects Agency, Arlington, VA. 2015. [Online]. Available: <https://www.darpa.mil/program/extreme-ddos-defense>
11. D. McPherson, D. Oran, D. Thaler, and E. Osterweil, "Architectural considerations of IP anycast," RFC-7094, 2014.
12. G. M. Moura et al., "Anycast vs. DDoS: Evaluating the November 2015 root DNS event," in *Proc. Internet Measurement Conf.*, ACM, 2016, pp. 255–270. doi: 10.1145/2987443.2987446.
13. "[state of the internet]/security: A year in review," Akamai, Cambridge, MA, vol. 4, no. 5, 2018.
14. T. Samson, "Spamhaus DDoS attack just another day for ISPs," *InfoWorld*, 2013. [Online]. Available: <https://www.infoworld.com/article/2613993/spamhaus-ddos-attack-just-another-day-for-isps.html>
15. "Distributed denial of service defense (DDoSD)," DHS Science and Technology Directorate, Washington, D.C., 2016. [Online]. Available: <https://www.dhs.gov/sites/default/files/publications/FactSheet%20DDoSD%20FINAL%20508%20OCC%20Cleared.pdf>
16. C. Morales, "NETSCOUT Arbor confirms 1.7 Tbps DDoS attack; The terabit attack era is upon us," NETSCOUT: Westford, MA, Mar. 5, 2018. [Online]. Available: <https://www.netscout.com/blog/asert/netscout-arbor-confirms-17-tbps-ddos-attack-terabit-attack-era>
17. E. Leverett and A. Kaplan, "Towards estimating the untapped potential: A global malicious DDoS mean capacity estimate," *J. Cyber Policy*, vol. 2, no. 2, pp. 195–208, 2017. doi: 10.1080/23738871.2017.1362020.
18. S. Mansfield-Devine, "The growth and evolution of DDoS," *Netw. Secur.*, vol. 2015, no. 10, pp. 13–20, 2015. doi: 10.1016/S1353-4858(15)30092-1.

**ERIC OSTERWEIL** is an assistant professor in the Department of Computer Science at George Mason University. Contact him at [eoster@gmu.edu](mailto:eoster@gmu.edu).

**ANGELOS STAVROU** is a professor in the Department of Computer Science at George Mason University. Contact him at [astavrou@gmu.edu](mailto:astavrou@gmu.edu).

**LIXIA ZHANG** is the Jonathan B. Postel Professor of Computer Science at the University of California, Los Angeles. Contact her at [lixia@cs.ucla.edu](mailto:lixia@cs.ucla.edu).



# A Collapsing Academy, Part 1

Hal Berghel, University of Nevada, Las Vegas

*There are ominous clouds on the academic horizon. I set forth some reasons below.*

**B**y the academy, I mean the collection of accredited undergraduate universities and research-oriented graduate programs. We need not put a fine point on this definition, for what I say below applies to virtually the entire academic frontier.

## UNHINGING THE ACADEMY FROM CORE PRINCIPLES

Most of the top-tier state universities are state supported in name only. That wasn't the case 50 years ago. But over the past half century, there has been a steady erosion of state financial support for public postsecondary education. This has been replaced by increases in student tuition and fees, federal support for specific initiatives (for example, the G.I. Bill and Title IX), charitable contributions, business support of special programs, cost-sharing revenue from external (grant) funding, and the like. At this point, most of the larger public universities, and all of the more prestigious public universities, receive less than half of their revenue from state coffers, and that revenue percentage decreases every year. A half century ago the term *public*

university meant that the primary support was tax dollars. But for many years, politicians have rejected the premise that support of public universities is a public responsibility. Many of you will remember that enrollment in the University of California system was tuition free until Ronald Reagan became governor. Any resident of the state who qualified for admission to the University of California system received a tuition waiver. Even if the tuition wasn't free in most other states, it was heavily subsidized. This isn't ancient history; it was just a few decades ago.

We will set aside the question of whether and to what extent taxpayer support of higher education is a public good. Instead, I'll deal with a less controversial issue: the negative consequences that follow from the erosion of public support. The most obvious downside is that the lack of tax support has increased the financial burden on the students through increases in tuition and fees. I remember a university president proclaiming that his tripling of tuition would not create a heavy burden for the students because he had identified a plentiful supply of private, high-interest loans. This overlooked two second-order downsides: 1) the George W. Bush administration changed the bankruptcy laws so that students could never get out from under their student debt (they were singled out as a special class of undeserving debtors in this regard), and 2) this is the same sort of de facto economic slavery that was

used against the sharecroppers after the Civil War. The idea that heaping permanent debt on students might offset the advantage of the plentiful supply of high-interest loans apparently never bothered him.

Contemporaneous with the shift of funding away from the taxpayer support were two other phenomena.

1. We entered the era of the professional administrator.
2. State and institutional leadership warmed to the concept of performance-based funding, although the phrase is actually a misnomer. Performance-based funding as it is applied in higher education circumscribes a family of metrics that purport to assess outcomes (nothing wrong with that), while in reality they just measure the academic beans that are the easiest to count.

Since the close connection between these two phenomena may not be obvious to nonacademics, I'll elaborate. In the gilded era of higher education, the 40 years after the end of the World War II, administrators tended to be drawn from the pool of faculty who were well respected for their academic prowess. In those years, administrators would not normally be confirmed by regents or trustees unless they were vetted by the faculty. These administrators 1) understood how a university worked; 2) subscribed to the core value of providing students with a diversified, well-rounded education; 3) recognized that the uniqueness of the every institution was one of its strengths; 4) were not disposed to mission creep; 5) had a strong commitment to the quality of the entire educational experience (which has degenerated into the goal of maximizing external funding); and 6) were relatively immune to both labor trends and academic fads. These values were then subsumed under a shared governance model that split the oversight between an administration and faculty.

This began to change about 50 years ago as we entered the era of the professional administrator (which coincided with the decline of the aforementioned gilded era). The professional administrators became increasingly distant from the core functions of the university (teaching, research, and service), less likely to be distinguished teacher-scholars themselves, and, as a result, less likely to enjoy the respect of the academic community they allegedly served. As the demand for faculty vetting diminished, other stakeholders like trustees, legislators, business leaders, and their lobbies, together with major benefactors, began to exert more control over the administrative selection process. And as their influence increased and the appreciation of institutional core goals decreased, a race to the bottom ensued, where efficiency and economy displaced the core academic principles discussed above. This is not to imply that efficiency and economy are necessarily at odds with lofty academic principles, but, as we shall see, the devil lay in the details.

This shift of emphasis from core academic principles to efficiency and cost cutting changed institutional priorities permanently.<sup>1,2</sup> It moved the modern taxpayer-supported university away from education, enlightenment, and literacy to indoctrination, skill development, and job training, while at the same time the shared governance model degenerated into a market-based free-for-all. Along the way, the quality of the senior academic leadership became less education- and student-centric and more expense minded, which produced a class of leadership that was the worst of two worlds: unqualified to run a profitable business and incapable of adding any real value to the educational experience. One major consequence of this duality was the race by professional administrators to performance-based funding—another one of those catchphrases that sounds good in principle but is in practice vacuous.

Under a performance-based funding model, units that underperform

will experience budget decline. This is the soft version of the “rank-and-yank” system that Enron used to become the prestigious corporate icon it is today. Under performance-based funding, everything rests on metrics. Since there is no way to measure the intangibles known as *quality*, *value*, or *public good*, the professional administrator-manager substitutes other measures that pretend to be their correlates. (Appearance is reality, after all.) We illustrate by means of the following commonly employed metrics:

- › the cost of a degree as measured on a per-student basis
- › the graduation rate measured as a percentage of students who follow their program of study through to graduation
- › enrollment per unit as measured by full-time equivalent or student credit hour
- › the academic progress rate as a percentage of total student body with a grade point average (GPA) > 2.0
- › the retention rate as a percentage of the current students who are retained in the program and proceed to the next term
- › the cost of class/program expressed as cost per student per class
- › the degrees in strategic areas expressed as a number or percentage of total degrees given [for example, science, technology, engineering, and mathematics (STEM)].

The first thing to notice is how easy these parameters are to measure (read: count). While there is no automated heuristic that can be used to assess educational quality, these metrics can all be measured perfectly well on a Commodore 80 spreadsheet. In this way, the performance-based funding model shifts the burden of thoughtful oversight and common sense to spreadsheet accounting.

We concede that all of these metrics seem reasonable on the surface.

It's only in the light of the practical consequences that the absurdity becomes transparent. It is through this pragmatic lens that any claim of social good disappears and the moral hazard rears its ugly head. (The moral hazard in this case incentivizes university employees to do those things that actually undercut their primary mission of delivering a quality education.)

This is certainly the case with metric 1, the cost of a degree, which, on the surface, seems to be a plausible measure of efficiency. After all, when it comes to public expense, less is better, right? However, this metric does not measure efficiency at all. For one thing, "cost" applies to infrastructural cost—not the cost of curriculum delivery. In addition to direct instructional expenses, institutional cost includes the amortized construction and maintenance expenses of buildings and grounds; the support of an entourage of assistant-, deputy associate-, and vice-presidents and chancellors in charge of virtually nothing important; institutional investments; facilities and administrative expenses; athletic program costs; and so forth, none of which are directly tied to the education of any student. Of course, classes must be held in buildings and clean bathrooms are required, but a very large percentage of an administrative budget for a university is dictated by legislative policies and administrative decisions independent of the expressed needs of faculty, students, and staff to deliver the curriculum—and may even be unknown to them.

Professional administrators, like their corporate counterparts, measure their importance in terms of the size of their budgets and not proof of whether anything important resulted from the expense. We note that in the calculation of metric 1 (TOTALBUDGET/#\_DEGREES), the professional administrators, legislators, and trustees control the numerator. If they want to claim increased efficiency, they either have to shrink the numerator (which would entail cutting their own budgets) or

grow the denominator. Talk about a no-brainer. Thus, academic units will be directed to increase the number of degrees if they want to protect their budgets. We note that no discussion of academic standards is involved. The professional administrator has steered the academy toward a diploma-mill model of productivity. Thus, metric 1 is not a useful measure of educational efficiency at all but simply a measure of administrative budget priorities. From a perspective of the faculty and academic units and academic standards, it is paradigmatically a moral hazard.

Metric 2, graduation rate, shares the same problem as metric 1. There are several factors that prevent graduation rate from being a reliable measure. First, based on my more than 40 years of experience in academia, the single most important factor in the failure of students to graduate is financial, so one consequence will be that institutions serving the more disadvantaged communities will find metric 2 the most onerous. As with cost of a degree, an academic unit that wants to avoid a budget penalty will be incentivized to increase the number of graduates in absolute terms and adjust matriculation standards accordingly. Other things being equal, academic standards are inversely related to graduation rates and enrollments. You can see where this is going. At any given time, there is a finite pool of qualified college applicants to go around, so if all schools draw more students from this finite pool, they will have to lower the admission standards to accommodate them all. This is what is informally known as the *butts in seats* dilemma. We note that this is exacerbated when students succumb to insurmountable financial pressure, which has the effect of further shrinking the pool of qualified students for reasons that have nothing to do with academic ability. Riddling such students with student debt may ameliorate the *butts in seats* problem for the institution, but it creates dire financial problems for students.

We note that metric 3, enrollment per unit, is a variation of the same theme shared with metrics 1 and 2. They all have counterproductivity and predictable adverse consequences in common. However, metric 4, academic progress rate, puts a different twist on the issue. While metrics 1–3 directly affect student enrollment, metric 4 affects student performance. The inevitable consequence of rewarding academic units that satisfy a proscribed overall GPA is grade inflation, pure and simple. When academic leadership dictates that failure to achieve a minimal overall student GPA will negatively impact the unit budget, the effect will be that the overall GPA will chase after the required minimum. (Crack addicts call this "chasing the bell.") Chairs and deans aren't stupid and know how to chase their proverbial budgetary bell.

The same holds true for metric 5, retention rate: if the academic unit is penalized for losing too many students to dropout and withdrawal, it will find creative ways to prevent the students from dropping out and withdrawing. To use coarse measures like unit GPA and student retention without addressing the underlying causes is absurd on its face. What is more, dropout and failure are as normal a part of education as they are in sports and business. Thus, it takes very little imagination to see that metrics 2–5 are direct contributors to a scholastic moral hazard—that is, the actual consequences are directly at odds with the very quality of educational delivery that the performance-based budgeting model promised to improve. The very fact that performance-based funding is taken seriously by legislators, trustees, and professional administrators shows that they are focused on the diploma and not the underlying quality of education. Why? Because diplomas are easy to count. Once again, all of these efforts are counterproductive in the sense that their effect is the opposite of their alleged intention. The same effect would result from any similar system



used in manufacturing if the budget of a quality control division was determined by ad hoc metrics that dealt only with the quantity of goods produced and cost per unit. Over time, the number of rejects and cost per unit will go down, and the output will increase. No news there. But any claim that these results are a useful measure of the quality of the goods produced is silly. In any competitive environment worthy of the name, a company that uses these metrics will fail.

We note that metric 6, cost of class/program, is a doubly bad metric as it fails to measure anything useful while also creating internal strife between academic units as they fight to avoid being at the bottom of the list of efficient programs. The inevitable consequence of this metric is huge classes where the class size works against the quality of instruction. Nowhere is this more problematic than in computer science and computer engineering, where the enrollments of critical courses that require extensive programming and interactivity have been driven to dangerous levels.

We conclude by addressing the politically hot-button issue of what constitutes a strategically important program of studies when this will likely be determined by the same academic leaders who control the overall cost of education as well as lobbying from corporate interests. One such strategic folly is motivated by the so-called “STEM crisis.” We’ll refer the reader to resources that the STEM crisis is now, always has been, and likely always will be, a myth that is propagated for the economic benefit of corporations to lower labor costs and leave it at that.<sup>3-6</sup>

## BAYH-DOLEFUL

While we’re on the subject of moral hazards, I’ll take the opportunity to discuss one of the most ill-advised pieces of federal legislation in the past 50 years, the Bayh-Dole Act (B-D).<sup>7</sup> This act was rushed through the 1980 lame duck session of Congress and is so hydra-headed that its ultimate effects

were virtually impossible to foresee. For present purposes, we narrow our attention to the single issue: whether any positive consequences of B-D could have been achieved by alternative legislation that avoided the negative, punitive tax consequences to the citizen. When approached from this perspective, it is not obvious that B-D was an overall public good. But, for better or worse, in terms of federally supported research, 1980 was a watershed year.

The alleged motivation of the bill was to facilitate technology transfer to the private sector. Proponents claimed that the federal government was spending billions of dollars on research that was not translating into commercialized products (that is, products that private industry could turn into a profit). We’ll pass over the fact that this claim was largely false.<sup>8,9</sup> But even if it were true, it didn’t logically follow that B-D was the most desirable path of legislative action. This observation was made at the time the bill was introduced and partially explains why support in Congress was slow to develop. In the end, the majority accepted without proof the claim that the recommended changes in U.S. patent policy could be leveraged to stir innovation and make the U.S. economy more competitive without significant expense to the taxpayer (which is also false). The accelerated congressional decision making is best understood in terms of a prevailing political attitude that any effort that would make private industry more profitable was desirable. The received view in Congress was that the focus should remain on innovation to the exclusion of any negative externalities like wealth transfer from taxpayer to corporations. The question of whether the legislation would be economically fair to the taxpayer was not taken seriously.

One of the subordinate claims by private-sector supporters was that future innovation and economic security demanded that the current federal patent policy be overturned. For one thing, existing policy specifically

prevented the exclusive licensing of federal patents. Corporatists argued that such restrictions were hostile to private enterprise. Specifically, the source of the hostility was thought to be two patent policies established in 1941: the policy known as *license model*, whereby the government retained a royalty-free license to use any federal patent, even if the license was sold to a private party, and the policy known as *title model*, where all federal licenses had to be non-exclusive.<sup>9-11</sup> Both models, the business lobby claimed, inhibited the stimulation of innovation by forcing competition on the licensees. Apparently, it was believed that anything less than unfair competitive advantage would stifle technology transfer.

While this aspect of the status quo had to be eliminated, a second aspect of the status quo was deemed absolutely essential: the federal subsidy of the costs of research. Bayh-Dole legislated that the federal taxpayer should continue to pay for research but give up any existing entitlement to recoupment and royalty-free license sharing. A blind eye to monopolistic practices that might result from exclusive licensing was also called for.

It is in this more complete context that B-D can be understood. In this context, it becomes clear that the common view that B-D has been an unqualified success as the driving force behind innovation in the United States for the past 40 years<sup>12</sup> is excessively simplistic. There is no denying that B-D made it easier for commercial interests to take advantage of patents resulting from federal research support. Nor can it be denied that universities and research centers that participated in the research have been able to derive considerable revenue from the patents that resulted. But in all other important areas, the public value of B-D is mixed at best. For example, one special provision in the original legislation restricted the exclusive licensing arrangements to small businesses, which has some appeal to the fair-minded set. However, this provision was disingenuous and

only lasted a few years. Ronald Reagan repealed this provision by executive order seven years later, so that even the largest corporations could exploit the monopolistic value of the exclusive licenses.<sup>13</sup> Similarly, recoupment and royalty-free use provisions in the original draft were equally disingenuous and quickly eliminated from consideration. But the provision that made it all of the way from first draft to final passage was the massive wealth transfer from the taxpayer to the private sector. Not surprisingly, the three groups who remained the strongest champions of B-D throughout were university research administrators, the business community, and legislators who supported such wealth transfers to the private sector on principle. The former two groups were motivated by economics, while the latter was driven by ideology.

We cannot accept the premise that B-D contributed to the velocity of innovation without also discussing the considerable cost to the taxpayer. The most important question at that time, and that remains today, is not whether B-D increased the commercialization of federally supported research but, rather, whether the benefits outweighed the costs. Specifically, intelligent analysis demands that we inquire whether alternative modifications of federal patent policy might have achieved much the same results without heaping such abuse on the federal taxpayer. This question is almost entirely overlooked by commentators, even those who claim objectivity in their reports to Congress.<sup>12</sup>

B-D perverts a pure capitalist model of risk management whereby the investor uses the best, but admittedly imperfect, knowledge available to determine whether a prediction of future sales of a product or service will both cover the costs of production and deliver a reasonable profit. In the simplest case, costs to commercialize patents arise from research and development. But under B-D, research costs of affected projects are subsidized and thus artificially low. However—and this is a

critical point—under B-D, the party that underwrites the research is not allowed to participate directly in any of the profit. What is more, if the license leads to a useful consumable, B-D guarantees that the taxpayer will pay even more, as he will have to also pay the licensee a profit. In plain terms, B-D guarantees that a taxpayer's price for any commercialized product of tax-supported research will always be inflated when compared to the pre-B-D federal patent policy. That is a logical consequence of the bill.

Thus, B-D corrupts the expected correlation of risk and reward. Of course, there are other models that are corrupting. Cost-plus contracts, for example, virtually eliminate risk. But at least they have the saving grace of having a cap put on profits. B-D minimizes an important part of the risk but without any corresponding limit on profit. Thus, unlike cost-plus contracts, B-D not only decouples risk and reward, it also introduces an asymmetry between risk and profit. In fact, from the taxpayer's point of view, B-D licensing actually creates an inverse relationship between the risk and profit. This is crony capitalism at its finest. To paraphrase country artist Jerry Reed, the corporations and universities got the mine, and the taxpayer got the shaft.

Of course, other observations may be made with regard to such issues as whether the absence of competition baked into B-D will lead to optimal allocation of resources. Economist Kenneth Arrow observed that there is a natural inclination for businesses to underinvest in research because of risk. In addition to this natural downward bias, there will also be a natural inclination to oppose any undertaking that does not lead to monopoly through exclusive licensing. But, Arrow argues, these conditions will ultimately “reduce the efficiency of inventive activity in general and will therefore reduce its quantity also.” Further, optimal technology transfer decisions will result from the least-restrictive flow

of research information even when the profit potential for any particular licensee may be suboptimal. Arrow offers a proof that the incentives to invent are greater in competitive markets in his 1962 paper.<sup>14</sup>

So, B-D actually creates its moral hazard by encouraging business behavior that 1) is unfair to a primary sponsor (taxpayer), 2) disincentivizes corporations to make optimal technology transfer decisions, 3) reduces the efficiency of inventive activity, and 4) makes the resulting markets less competitive. The B-D Act is a poster child for ill-advised legislation.

We repeat that these points were made as the legislation was introduced by public figures from Ralph Nader to Admiral Hyman Rickover.<sup>8,11,9</sup> But the appeal of reducing corporate risk while maximizing profit potential proved too powerful to overcome in Congress. Even research from a Nobel Laureate in economics didn't affect the deliberation. However, from the point of view of the country and the liberal politicians who were tricked into supporting it, the B-D Act has proven to be a Faustian bargain.

**W**e have given two examples of radical changes that have negatively impacted higher education: the move to professional administrators and the crony capitalistic way that B-D handled federally supported research. These changes weren't inevitable. In both cases, major negative consequences were anticipated by thoughtful scholars at the time these decisions were made. I have elsewhere used the term *spinfluenza* to describe the speed with which really bad ideas take hold over politicians, business, and administrative leaders.<sup>15</sup> At this point, spinfluenza in higher education has achieved pandemic proportions.

Those who are sympathetic to my arguments might ask how these mistakes might be undone. Obvious solutions are unrealistic. Professional administrators will not easily welcome

additional accountability for defending their budgets to legislators and trustees when the problem can be so easily offloaded to subordinate units. In this case, the correction would have to be top-down and inspired by state leadership. B-D is an entirely different can of worms. Perhaps the most direct approach would involve repeal. However, that would carry with it political liabilities from the donor class. However, much improvement could be achieved through the simple restoration of recoupment and nonexclusive licensing provisions. The strong suit of this approach is that the original arguments to repeal these provisions were so lame that they might be politically embarrassing to defend anew.

It must be admitted that what we've called the collapse of the academy is the gradual result of complex political and social forces that have surfaced in many factors and forms. A complete discussion would include the following (to name but a few)

- › the effect of changing the public's perception of faculty tenure
- › the dilution of shared governance and the subtle progression in the direction of authoritarianism
- › the thorny issue of what constitutes acceptable academic free speech
- › academic standards that have become moving targets
- › the question of how one might meaningfully measure quality scholarship
- › changing public expectations of postsecondary education
- › the rise in importance of narrow-focus stakeholders and their effect on institutional decision making
- › the widespread acceptance of donations and gifts that are restricted to uses that support particular ideologies and belief systems
- › the political antagonism to the principle of a diversified,

well-rounded education by groups who seek to maximize uniformity of beliefs and consent

- › the impact of social media on the educational experience
- › the pressure for online delivery to reduce the cost of service.

We will cover some of these topics in future columns. Throughout this series, we advance the notion that the race to academic postmodernity is inconsistent with those academic principles that led us to the economic success and quality of life that we currently enjoy. ■

## REFERENCES

1. B. Ginsberg, *The Fall of the Faculty: The Rise of the All-Administrative University and Why It Matters*. New York: Oxford Univ. Press, 2011.
2. P. Eckel, *Changing Course: Making the Hard Decisions to Eliminate Academic Programs* (Studies on Higher Education), 2nd ed. Lanham, MD: Rowman and Littlefield, 2009.
3. H. Berghel, "STEM crazy," *Computer*, vol. 48, no. 9, pp. 75–80, Sept. 2015. doi: 10.1109/MC.2015.256.
4. H. Berghel, "STEM revisited," *Computer*, vol. 47, no. 3, pp. 70–73, Mar. 2014. doi: 10.1109/MC.2014.72.
5. R. N. Charette, "The STEM crisis is a myth," *IEEE Spectrum*, Aug. 30, 2013. Accessed on: Apr. 26, 2020. [Online]. Available: <http://spectrum.ieee.org/at-work/education/the-stem-crisis-is-a-myth>
6. M. S. Teitelbaum, *Falling Behind? Boom, Bust, and the Global Race for Scientific Talent*. Princeton, NJ: Princeton Univ. Press, 2014.
7. U.S. Senate and House. 96th Congress. (1980, Dec. 12). Public Law 96-517, Chapter 30: Prior art citations to office and reexamination of patents, (aka the Bayh-Dole Act), 94 STAT. 3015. Accessed on: Apr. 26, 2020. [Online]. Available: <https://history.nih.gov/research/downloads/PL96-517.pdf>
8. M. Kenney and D. Patton, "Reconsidering the Bayh-Dole Act and the current university invention ownership model," *Res. Policy*, vol. 38, no. 9, pp. 1407–1422, Nov. 2009. doi: 10.1016/j.respol.2009.07.007.
9. J. Washburn, *University Inc.: The Corporate Corruption of Higher Education*, New York: Basic Books, 2005.
10. R. S. Eisenberg, "Public research and private development: Patents and technology transfer in government-sponsored research," *Virginia Law Rev.*, vol. 82, no. 8, pp. 1663–1727, 1996. doi: 10.2307/1073686.
11. R. Nelson, "The market economy, and the scientific commons," *Res. Policy*, vol. 33, no. 3, pp. 455–471, Apr. 2004. doi: 10.1016/j.respol.2003.09.008.
12. W. H. Schacht, "The Bayh-Dole Act: Selected issues in patent policy and the commercialization," Congressional Research Service (CRS) Report for Congress prepared for members and committees of Congress, Rep. no. RL32076, Dec. 3, 2012. Accessed on: Apr. 26, 2020. [Online]. Available: <https://fas.org/sgp/crs/misc/RL32076.pdf>
13. Exec. Order No. 12,591, R. Reagan. (1987, Apr. 10). *Facilitating Access to Science and Technology*, cf. 1(b)(4). Accessed on: Apr. 26, 2020. [Online]. Available: [https://ipmall.law.unh.edu/sites/default/files/BAYHDOLE/latkinPDF/Executive\\_Order\\_No.\\_12591\\_Office\\_of\\_the\\_Press\\_Secretary\\_4-10-1987.pdf](https://ipmall.law.unh.edu/sites/default/files/BAYHDOLE/latkinPDF/Executive_Order_No._12591_Office_of_the_Press_Secretary_4-10-1987.pdf)
14. K. Arrow, "Economic welfare and the allocation of resources for invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, National Bureau of Economic Research. Princeton, NJ: Princeton Univ. Press, 1962, pp. 609–626.
15. H. Berghel, "Borderline executive disorder," *Computer*, vol. 48, pp. 82–86, Apr. 2015. doi: 10.1109/MC.2015.96.

**HAL BERGHEL** is a Fellow of the IEEE and ACM and a professor of computer science at the University of Nevada, Las Vegas. Contact him at [h1b@computer.org](mailto:h1b@computer.org).



# HOST 2020

6–9 Dec. 2020 • San Jose, CA

REGISTER NOW!



## IEEE INTERNATIONAL SYMPOSIUM ON HARDWARE-ORIENTED SECURITY AND TRUST

6–9 Dec. 2020 • San Jose, CA, USA • DoubleTree by Hilton

Join dedicated professionals at the IEEE International Symposium on Hardware Oriented Security and Trust (HOST) for an in-depth look into hardware-based security research and development.

### Key Topics:

- Semiconductor design, test and failure analysis
- Computer architecture
- Systems security
- Cryptography and cryptanalysis
- Imaging and microscopy

Discover innovations from outside your sphere of influence at HOST. Learn about new research that is critical to your future projects. Meet face-to-face with researchers and experts for inspiration, solutions, and practical ideas you can put to use immediately.

**REGISTER NOW:** [www.hostsymposium.org](http://www.hostsymposium.org)





## IEEE INTELLIGENT SYSTEMS “AI’S 10 TO WATCH” AWARD: CALL FOR NOMINATIONS

IEEE Intelligent Systems solicits nominations for its 2020 “AI’s 10 to Watch” Award. Candidates must have received their Ph.D.s in 2014 or later and have already made significant contributions to one or more areas of artificial intelligence (AI).

The selection committee is chaired by a member of the IEEE Intelligent Systems Editorial Board and consists of prominent AI researchers from various AI subfields and geographic regions of the world. To nominate a candidate, please provide the following (preferably as a zip file):

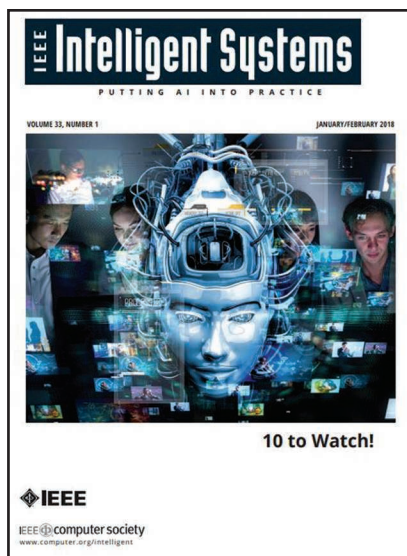
- › a CV, including a publication list
- › two reference letters
- › a statement of not more than 200 words that summarizes the candidate’s achievements in AI.

Nominations should be sent to: [ieee@ten-to-watch-in-ai.com](mailto:ieee@ten-to-watch-in-ai.com). The deadline is 10 July 2020. View past awardees at [www.computer.org/press-room/2018-news/ieee-intelligent-systems-ai-10-to-watch](http://www.computer.org/press-room/2018-news/ieee-intelligent-systems-ai-10-to-watch) and <http://doi.ieeecomputersociety.org/10.1109/MIS.2018.012001549>.

## CALL FOR IEEE COMPUTER SOCIETY AWARDS NOMINATIONS

The following IEEE Computer Society (IEEE CS) awards are seeking nominations by 1 October 2020:

- › Taylor L. Booth Education Award: [www.computer.org/volunteering/awards/booth](http://www.computer.org/volunteering/awards/booth)
- › Mary Kenneth Keller Computer Science and Engineering Undergraduate Teaching Award: [www.computer.org/volunteering/awards/cse-undergrad-teaching](http://www.computer.org/volunteering/awards/cse-undergrad-teaching)



- › Charles Babbage Award: [www.computer.org/volunteering/awards/babbage](http://www.computer.org/volunteering/awards/babbage)
- › Computer Entrepreneur Award: [www.computer.org/volunteering/awards/entrepreneur](http://www.computer.org/volunteering/awards/entrepreneur)
- › Edward J. McCluskey Technical Achievement Award: [www.computer.org/volunteering/awards/technical-achievement](http://www.computer.org/volunteering/awards/technical-achievement)
- › Harry H. Goode Memorial Award: [www.computer.org/volunteering/awards/goode](http://www.computer.org/volunteering/awards/goode)
- › Hans Karlsson Standards Award: [www.computer.org/volunteering/awards/karlsson](http://www.computer.org/volunteering/awards/karlsson)
- › Richard E. Merwin Award for Distinguished Service: [www.computer.org/volunteering/awards/merwin](http://www.computer.org/volunteering/awards/merwin)
- › Women of the ENIAC Computer Pioneer Award: [www.computer.org/volunteering/awards/pioneer](http://www.computer.org/volunteering/awards/pioneer)
- › W. Wallace McDowell Award: [www.computer.org/volunteering/awards/mcdowell](http://www.computer.org/volunteering/awards/mcdowell)
- › Harlan D. Mills Award: [www.computer.org/volunteering/awards/mills](http://www.computer.org/volunteering/awards/mills)

The following IEEE CS Volunteer Service Awards accept nominations throughout the year:

- › T. Michael Elliott Distinguished Service Certificate: [www.computer.org/volunteering/awards/distinguished-service](http://www.computer.org/volunteering/awards/distinguished-service)
- › Meritorious Service Certificate: [www.computer.org/volunteering/awards/meritorious-service](http://www.computer.org/volunteering/awards/meritorious-service)
- › Outstanding Contribution Certificate: [www.computer.org/volunteering/awards/outstanding-contribution](http://www.computer.org/volunteering/awards/outstanding-contribution)
- › Continuous Service Certificate: [www.computer.org/volunteering/awards/continuous-service](http://www.computer.org/volunteering/awards/continuous-service)
- › Certificate of Appreciation: [www.computer.org/volunteering/awards/certificate-of-appreciation](http://www.computer.org/volunteering/awards/certificate-of-appreciation)

## AHMED LOURI TO RECEIVE 2020 IEEE COMPUTER SOCIETY EDWARD J. MCCLUSKEY TECHNICAL ACHIEVEMENT AWARD



Ahmed Louri

Ahmed Louri, an IEEE Fellow, will receive the 2020 IEEE CS Edward J. McCluskey Technical Achievement Award “for pioneering contributions to the solu-

tion of on-chip and off-chip communication problems for parallel computing and manycore architectures.”

The Edward J. McCluskey Technical Achievement Award is given for outstanding and innovative contributions to the fields of computer and information science and engineering or computer technology, usually within the past 10–15 years. Contributions must have significantly promoted technical progress in the field. The award consists of a certificate and a US\$2,000 honorarium.

Louri is the David and Marilyn Karlgaard Endowed Chair Professor of Electrical and Computer Engineering at George Washington University (GW). He is the director of GW's High-Performance Computing Architectures and Technologies Laboratory.

For more than 30 years, Louri has done seminal work on the development of general frameworks for scalable, energy-efficient, high-performance, and reliable communications for parallel computing systems. In particular, he has accomplished fundamental and pioneering work in the critical areas of scalable and reliable multiprocessors and energy-efficient, fault-tolerant, and high-performance network-on-chips for multicore and manycore architectures. His recent research focuses on accelerator-rich reconfigurable heterogeneous architectures; machine learning techniques for efficient computing, memory, and interconnect systems; and future parallel computing models and architectures, including deep neural networks.

He has published more than 180 refereed journal articles and peer-reviewed conference papers. He is the co-inventor on several U.S. and international patents. Louri's research has been sponsored by the National Science Foundation (NSF); U.S. Department of Energy; Air Force Office of Scientific Research; and companies including Intel, IBM, Cisco, Oracle, Raytheon, Physical Optics Corporation, and US West Technologies. He has received best article awards and best paper awards at several conferences, and he has been named a distinguished visiting fellow at several international institutions.

Louri served as a program director in the NSF Directorate for Computer and Information Science and Engineering.

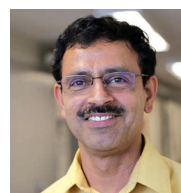
He directed the core computer architecture program and was on the management team of several cross-cutting programs. While at the NSF, he initiated multidisciplinary research programs in several key areas of computer architecture, high-performance computing, emerging technologies, resiliency, and security. The most notable among these is the Scalable Parallelism in the Extreme Program (SPX), which is now in its sixth year.

Louri served as steering committee chair, program committee chair, general chair, and keynote speaker of many IEEE/ACM international conferences and symposia, including general chair for the IEEE International Symposium on High-Performance Computer

Architecture (HPCA) 2019, HPCA 2007, and the 2020 IEEE International Conference on High-Performance Computing and Communications and the program committee area chair of the 2020 IEEE International Parallel and Distributed Processing Symposium. He is the editor-in-chief of *IEEE Transactions on Computers*, the flagship journal for the IEEE CS, and also serves as an associate editor for *IEEE Transactions on Sustainable Computing* and *IEEE Transactions on Cloud Computing*.

More information about the award can be found at [www.computer.org/volunteering/awards/technical-achievement](http://www.computer.org/volunteering/awards/technical-achievement).

## **B.S. MANJUNATH TO RECEIVE 2020 IEEE COMPUTER SOCIETY EDWARD J. MCCLUSKEY TECHNICAL ACHIEVEMENT AWARD**



B.S. Manjunath

B.S. Manjunath, an IEEE Fellow, will receive the 2020 IEEE CS Edward J. McCluskey Technical Achievement Award "for contributions to

image search, retrieval, and bio-image informatics."

The Edward J. McCluskey Technical Achievement Award is given for outstanding and innovative contributions to the fields of computer and information science and engineering or

**The Edward J. McCluskey Technical Achievement Award is given for outstanding and innovative contributions to the fields of computer and information science and engineering or computer technology.**

computer technology, usually within the past 10–15 years. Contributions must have significantly promoted technical progress in the field. The award consists of a certificate and a US\$2,000 honorarium.


Manjunath is a distinguished professor of electrical and computer engineering and directs the Center for Multimodal Big Data Science and Healthcare at the University of California, Santa Barbara. He is highly acclaimed for technical excellence in bringing computer vision to a diversity of applications that range from digital libraries/remote sensing to marine sciences, materials science, and biology and for

the development of the reproducible scientific image analytics platform BisQue. He has made fundamental contributions to image/video segmentation, image forensics, registration, feature extraction, image search, and retrieval that have been documented in more than 300 peer-reviewed publications. The impact of his technical work is demonstrated by an h-index of 70 (Google Scholar) and more than 33,000 citations, with six of his publications cited more than 1,000 times. He holds 24 U.S. and international patents, many of which relate to technologies that were

incorporated into the ISO/MPEG-7 standard. BisQue software is now adopted as a core service by the NSF-supported CyVerse Cyberinfrastructure (<http://cyverse.org>) that serves thousands of active users.

Manjunath is a coauthor of a paper that received the 2013 IEEE Transactions on Multimedia Best Paper Award and a coauthor of a paper that received the Excellent Paper Award at the 2013 ACM International Conference on Distributed Smart Cameras. He is a successful entrepreneur and has transitioned technologies relating to visual search,

computer security, and media forensics applications through Mayachitra, Inc., which he cofounded.

More information about the award can be found at [www.computer.org/volunteering/awards/technical-achievement](http://www.computer.org/volunteering/awards/technical-achievement). 

#### COMING UP NEXT

The IEEE CS election materials will be featured in the August issue of *Computer*.



**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. **OMBUDSMAN:** Email [ombudsman@computer.org](mailto:ombudsman@computer.org)  
**COMPUTER SOCIETY WEBSITE:** [www.computer.org](http://www.computer.org)

#### EXECUTIVE COMMITTEE

**President:** Leila De Floriani; **President-Elect:** Forrest Shull; **Past President:** Cecilia Metra; **First VP:** Riccardo Mariani; **Second VP:** Sy-Yen Kuo; **Secretary:** Dimitrios Serpanos; **Treasurer:** David Lomet; **VP, Membership & Geographic Activities:** Yervant Zorian; **VP, Professional & Educational Activities:** Sy-Yen Kuo; **VP, Publications:** Fabrizio Lombardi; **VP, Standards Activities:** Riccardo Mariani; **VP, Technical & Conference Activities:** William D. Gropp; **2019-2020 IEEE Division VIII Director:** Elizabeth L. Burd; **2020-2021 IEEE Division V Director:** Thomas M. Conte; **2020 IEEE Division VIII Director-Elect:** Christina M. Schober

#### BOARD OF GOVERNORS

**Term Expiring 2020:** Andy T. Chen, John D. Johnson, Sy-Yen Kuo, David Lomet, Dimitrios Serpanos, Forrest Shull, Hayato Yamana  
**Term Expiring 2021:** M. Brian Blake, Fred Douglass, Carlos E. Jimenez-Gomez, Ramalatha Marimuthu, Erik Jan Marinissen, Kunio Uchiyama  
**Term Expiring 2022:** Nils Aschenbruck, Ernesto Cuadros-Vargas, David S. Ebert, William Gropp, Grace Lewis, Stefano Zanero

#### BOARD OF GOVERNORS MEETING

24 – 25 September 2020 in McLean, Virginia, USA

#### 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 Tubby; **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](mailto:help@computer.org)

**Los Alamitos:** 10662 Los Vaqueros Cir., Los Alamitos, CA 90720; **Phone:** +1 714 821 8380; **Email:** [help@computer.org](mailto:help@computer.org)

**MEMBERSHIP & PUBLICATION ORDERS:** **Phone:** +1 800 678 4333; **Fax:** +1 714 821 4641; **Email:** [help@computer.org](mailto:help@computer.org)

#### IEEE BOARD OF DIRECTORS

**President & CEO:** Toshio Fukuda  
**President-Elect:** Susan K. "Kathy" Land  
**Past President:** José M.F. Moura  
**Secretary:** Kathleen A. Kramer  
**Treasurer:** Joseph V. Lillie  
**Director & President, IEEE-USA:** Jim Conrad; **Director & President, Standards Association:** Robert S. Fish; **Director & VP, Educational Activities:** Stephen Phillips; **Director & VP, Membership and Geographic Activities:** Kukjin Chun; **Director & VP, Publication Services & Products:** Tapan Sarkar; **Director & VP, Technical Activities:** Kazuhiro Kosuge



revised 1 May 2020





# CALL FOR SPECIAL ISSUE PROPOSALS

*Computer* solicits special issue proposals from lead experts. Proposed themes/issues should address timely, emerging topics that will be of broad interest to *Computer*'s readership. Special issues are an important component of *Computer*, as they deliver essential research insights and well-developed perspectives on new and established technologies and computing strategies.

We encourage submissions of high-quality proposals for the 2021 editorial calendar. Of particular interest are proposals centered on:

- offsite educational and business continuity technology challenges,
- privacy related to personal location tracking and surveillance (digital and physical),
- artificial intelligence and machine learning,
- technology's role in disrupted supply chains,
- misinformation and disinformation (fake information—malicious or non-malicious), and
- cyberwarfare/cyberterrorism

Proposal guidelines are available at:

[www.computer.org/csdl/magazine/co/write-for-us/15911](http://www.computer.org/csdl/magazine/co/write-for-us/15911)

Deadline for proposal submission: 15 September 2020



IEEE  
COMPUTER  
SOCIETY



IEEE



# Get Published in the New *IEEE Open Journal of the Computer Society*

**Submit a paper today to the premier new open access journal in computing and information technology.**

Your research will benefit from the IEEE marketing launch and 5 million unique monthly users of the IEEE *Xplore*® Digital Library. Plus, this journal is fully open and compliant with funder mandates, including Plan S.

**Submit your paper today!**

Visit [www.computer.org/oj](http://www.computer.org/oj) to learn more.

