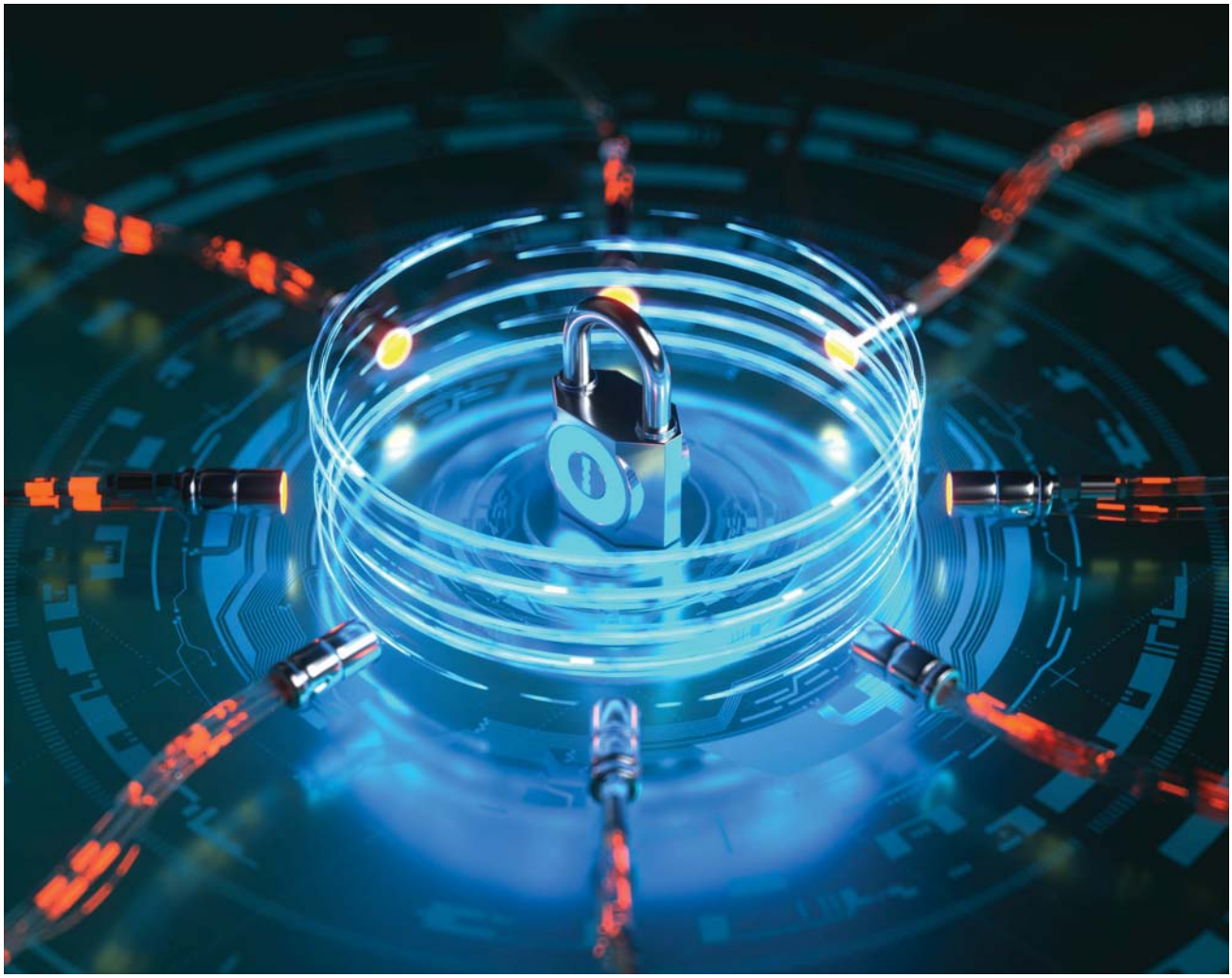
IT Professional
Technology Solutions for the Enterprise

VOLUME 20, NUMBER 3 MAY/JUNE 2018



Cyberthreats and Security

IEEE computer society
www.computer.org/itpro

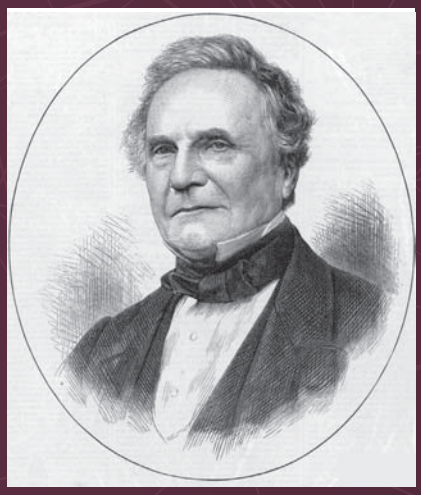
2019

IEEE-CS
Charles Babbage Award

CALL FOR AWARD NOMINATIONS
Deadline 1 October 2018

ABOUT THE IEEE-CS CHARLES BABBAGE AWARD
Established in memory of Charles Babbage in recognition of significant contributions in the field of parallel computation. The candidate would have made an outstanding, innovative contribution or contributions to parallel computation. It is hoped, but not required, that the winner will have also contributed to the parallel computation community through teaching, mentoring, or community service.

ABOUT CHARLES BABBAGE
Charles Babbage, an English mathematician, philosopher, inventor and mechanical engineer who is best remembered now for originating the concept of a programmable computer.



CRITERIA
This award covers all aspects of parallel computing including computational aspects, novel applications, parallel algorithms, theory of parallel computation, parallel computing technologies, among others.

AWARD & PRESENTATION
A certificate and a $1,000 honorarium presented to a single recipient. The winner will be invited to present a paper and/or presentation at the annual IEEE-CS International Parallel and Distributed Processing Symposium (IPDPS).

NOMINATION SUBMISSION
Open to all. Nominations are being accepted electronically at www.computer.org/web/awards/charles-babbage. Three endorsements are required. The award shall be presented to a single recipient.

NOMINATION SITE
awards.computer.org

AWARDS HOMEPAGE
www.computer.org/awards

CONTACT US
awards@computer.org

IT Professional

May/June 2018
Vol. 20, No. 3

www.computer.org/itpro

TABLE OF CONTENTS

Also in This Issue

ISSN: 1520-9202 Published by the IEEE Computer Society

IT Professional

EDITORIAL OFFICE

Publications Coordinator: itpro-ma@computer.org
Authors: www.computer.org/itpro/author.htm
Letters to the editors: itpro@computer.org
Subscribe: www.computer.org/subscribe
Change of address: address.change@ieee.org
Missing or damaged copies: help@computer.org
Reprints of articles: itpro-ma@computer.org
IT PROFESSIONAL
c/o IEEE Computer Society
10662 Los Vaqueros Circle, Los Alamitos, CA 90720 USA
Phone +1 714 821 8380; Fax +1 714 821 4010
www.computer.org/it-professional

Circulation: *IT Professional* (ISSN 1520-9202) is published bimonthly 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, PO Box 3014, Los Alamitos, CA 90720-1314; voice +714 821 8380; fax +714 821 4010; IEEE Computer Society Headquarters, 1828 L St. NW, Suite 1202, Washington, DC 20036. Visit www.computer.org/subscribe for subscription information. **Postmaster:** Send undelivered copies and address changes to IT Professional, Membership Processing Dept., IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854-4141. Periodicals Postage Paid at New York, NY, and at additional mailing offices. Canadian GST #125634188. Canada Post Publications Mail Agreement Number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8, Canada. Printed in the USA. **Editorial:** Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *IT Professional* does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space. IEEE prohibits discrimination, harassment, and bullying: For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

Cyberthreats and Security

Morris Chang
University of South Florida

Rick Kuhn
NIST

Tim Weil
Alcohol Monitoring Systems

One of the most challenging aspects of cybersecurity is that the problem space grows larger every year as more and more of everyday life is converted to digital activity. It is hard to think of any aspect of life today that does not involve IT for most of the population. Socializing, banking, shopping, dating, and healthcare are all done at least in part online. The potential for privacy violations and security challenges is seen in daily news reports. As an example of everyday cyberthreat and security protection, by the time this issue goes to press, the EU's General Data Protection Regulation (GDPR) will have gone into effect. Will this industry mandate improve online privacy protection by making the reporting of data breaches a mandatory requirement for international commerce? Or will more phishing and social engineering attacks take advantage of GDPR policies?

Cyberthreats should not be thought of just in the context of IT security and privacy design. Adequate cybersecurity must involve the active participation of everyone in an organization, as well as users. Although this can be seen as an enormous burden, the nature of technology is such that humans have been responding to challenges and adapting to complex environments for millennia, as well as systematizing solutions for particular applications. Approaches generally reflect some variation on the common-sense method of evaluating the problem, preparing, acting, and assessing the results.

Managers learn a *plan-do-check-act* cycle. Fighter pilots are taught to *observe-orient-decide-act*. In cybersecurity, the latest incarnation of this common-sense approach is the popular NIST Cybersecurity Framework, which teaches *identify-protect-detect-respond-recover*. As in other fields, these activities are intended to be performed in a continuous cycle, modifying plans and actions as the organization learns from successes and failures.

This issue includes articles that touch on all of the activities described above, and in some cases more than one phase of the Cybersecurity Framework cycle. We leave it as an exercise for the reader to decide how the lessons of each article fit into the different phases of the cycle.

As cybersecurity is involved with nearly all aspects of life, it is not possible to cover all types of security challenges in detail. Instead, the articles describe novel and interesting techniques that promote creative ways of thinking about cybersecurity in a broad range of applications.

IN THIS ISSUE

In "Advancing Cybersecurity: The Growing Need for a Cyber-Resiliency Workforce," authors Logan O. Mailloux and Michael R. Grimaila address the topic of preparing the next generation of cybersecurity professionals who must focus on cyber resiliency—bouncing back from computing faults, networking failures, cyberattacks, and unpredictable events—especially as the world becomes more connected via cyber-physical systems. They uniquely detail several key responsibilities, work roles, and expertise areas for the future cyber-resiliency workforce.

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

"The New Threats of Information Hiding: The Road Ahead" by Krzysztof Cabaj, Luca Caviglione, Wojciech Mazurczyk, Steffen Wendzel, Alan Woodward, and Sebastian Zander deals with the threat of using steganography (information hiding) to empower malware. The authors provide an overview of information-hiding techniques that can be utilized by malicious software, showcase existing and emerging threats, and discuss the future research directions to circumvent such threats. Industries and governments are currently asking these questions, and IT professionals should be aware of the issues involved.

In "Internet of Things Forensics: The Need, Process Models, and Open Issues," Maxim Chernyshev, Sherali Zeadally, Zubair Baig, and Andrew Woodward assess how the Internet of Things (IoT) paradigm brings a set of unique and complex challenges to the field of digital forensics. They provide a review of the state of the art of conceptual digital forensic models that can be applied to the IoT environment and discuss open issues that exist in these techniques when applied to IoT devices. This field is complex, in particular because of the security tradeoffs, but solutions apply to many other industries as well.

"Experiments with Ocular Biometric Datasets: A Practitioner's Guideline" by Zahid Akhtar, Gautam Kumar, Sambit Bakshi, and Hugo Proenca deals with ocular biometrics, where an individual is recognized via iris, retina, sclera, periocular region, or eye movements. This biometric trait is gaining more popularity in applications ranging from international border crossings to unlocking smart devices due to its ease of use and few user-cooperation requirements. The authors provide a review of ocular databases available in the literature, discuss diversities among these databases, and outline how to choose the proper database for experimentation.

In "The Evolving Cyberthreat to Privacy," A.J. Burns and Eric Johnson analyze breaches of personally identifiable information and find that they are significantly larger than other types of breaches. This shows that past breaches can be useful for predicting and mitigating future breaches. Considering the basic principles involved can spur creative thinking about how to improve cyber defenses.

We hope the articles in this issue will encourage readers to think about cybersecurity in new ways.

Finally, an article that was submitted as a general paper but fit the theme of this issue argues that despite the benefits of big data systems, they exhibit serious concerns for user privacy. In "Understanding Privacy Violations in Big Data Systems," Jawwad A. Shamsi and Muhammad Ali Khojaye provide an overview of privacy in the context of big data, categorizing four types of existing privacy violations in big data systems and suggesting countermeasures that can be taken. Although this article was not considered as part of the special issue and was accepted by other reviewers, we thought it important to include it in this special issue on cyberthreats and security.

We hope the articles in this issue will encourage readers to think about cybersecurity in new ways. Successfully addressing the cybersecurity needs of new technologies is not an easy task, but advances in data analytics, forensics, threat modeling, and other techniques presented in these articles can help us meet the challenge.

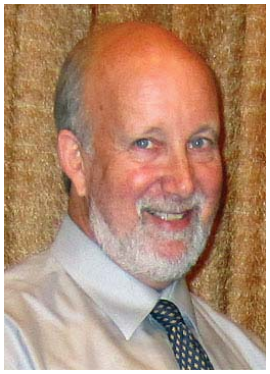
DISCLAIMER

Certain products may be identified in this document, but such identification doesn't imply recommendation by NIST or other agencies of the US Government, nor does it imply that the products identified are necessarily the best available for the purpose.

ABOUT THE AUTHORS

Morris Chang is a professor at the University of South Florida. Contact him at chang5@usf.edu.

Rick Kuhn is a computer scientist at NIST. Contact him at kuhn@nist.gov.



Tim Weil is a network project manager at Alcohol Monitoring Systems. Contact him at trweil@ieee.org.

Advancing Cybersecurity

The Growing Need for a Cyber-Resiliency Workforce

Logan O. Mailloux
Air Force Institute of Technology

Michael R. Grimaila
Air Force Institute of Technology

As the world becomes more dependent on connected cyber-physical systems, the cybersecurity workforce must adapt to meet these growing needs. The authors present the notion of a cyber-resiliency workforce to prepare the next generation of cybersecurity professionals.

The ever-growing demand for cyber-enabled systems and services has made cybersecurity one of the most serious challenges we face in the 21st century.[1] Moreover, the increasing complexity of modern safety-critical systems such as automobiles and commercial aircraft makes these advanced cyber-physical systems difficult to secure.[2] For example, the 2017 Ford F-150, a relatively common vehicle, runs software compiled from more than 150 million lines of source code.[3] Because these advanced cyber-physical systems are composed of multiple subsystems with special-purpose networks and dozens of processors, they are especially difficult to secure.[4]

Recent examples of such cyber-physical system vulnerabilities include the widely publicized *Wired* hacking demonstration against a Jeep Cherokee and the alleged compromise of a commercial airliner. In the first example, security researchers electronically controlled the Jeep from afar via remote connectivity features.[5] In the latter example, a single security researcher claimed to force the aircraft into an unplanned maneuver via the onboard entertainment system.[6]

Perhaps the best-documented open source example of cyber-physical system vulnerabilities is captured in Stephen Checkoway and his colleagues' research detailing several attack paths (shown in Figure 1).[7] This systematic work demonstrates the need for improved systems security approaches that include developing, operating, and maintaining systems designed and built to be resilient in the face of disruptive events and cyberattacks.

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

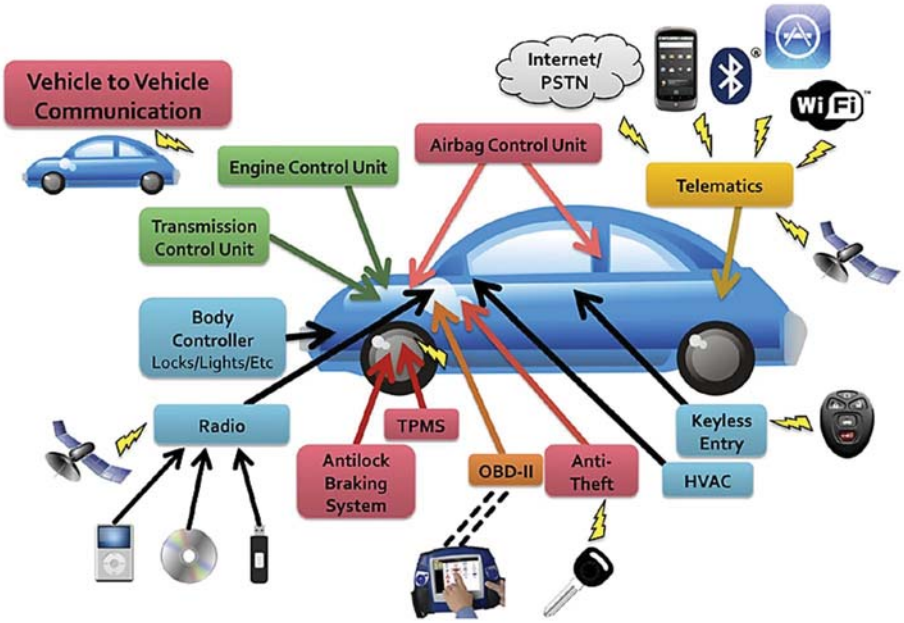


Figure 1. The many attack paths available against automotive vehicles.[7] (Reprinted with permission from S. Checkoway et al.)

UNDERSTANDING THE CYBER-RESILIENCY PROBLEM

The fact that cyber-physical systems require more cybersecurity attention has been brought to light through several initiatives across industry, academia, and even the US Department of Defense.[8] For example, last year the security training and certification organization SANS hosted their first Automotive Cybersecurity Training Summit. Likewise, a five-year collaboration between NIST, the NSA, and MITRE Corporation (with backing from several industry partners) culminated in the recent publication of NIST SP 800-160, which brings new life to the specialty discipline of systems security engineering.[9]

Although the science of "cyber resiliency"—the ability of a cyber-physical system to anticipate, withstand, and recover from adverse events—has been slow to develop, MITRE's Cyber Resiliency Engineering Framework provides a good starting point for discussion.[10] Before diving further into the topic, let's consider for a moment what cyber resiliency is and how it is different from cybersecurity.

At its essence, cyber resiliency focuses on bouncing back from computing faults, networking failures, cyberattacks, and unpredictable events.

At its essence, cyber resiliency focuses on bouncing back or fighting through computing faults, networking failures, cyberattacks, and unpredictable events.[11] More precisely, this means cyber-physical systems are required to maintain essential operational capabilities regardless of the threats they face (malicious or non-malicious) or where they originate (natural or man-made).[12] Thus, the resiliency of cyber-physical systems is much more than merely building more secure networks and software—it requires the ability to plan for and recover from hazardous events (expected or unexpected) through designing and building in affordable and trustworthy security and resiliency features.[13] Although many formal definitions of resiliency exist,[11] a readily understandable working definition of resiliency for cyber-physical systems is provided in Table 1.[10,14]

Table 1. Definition and attributes of resiliency for cyber-physical systems.

Term	Definition
Resiliency	The ability of a cyber-physical system to anticipate, withstand, and recover from adverse events.
Attribute	**Description**
Anticipate	Preparations for known, predicted, and unknown events to include changes in the operational environment, modes of operation, business/mission functions, emerging threats, integration of novel technologies, and other necessary changes.
Withstand	To absorb the negative impacts of adverse events such as system faults, user errors, software bugs, hardware failures, and cyberattacks.
Recover	To restore operations (and desired functionality) to an acceptable level within a specified time and performance requirement. Ideally, recovery also includes the ability of the system to "adapt" to reduce the impact of future adverse events.

Table 2 provides a comparison of why the security and resiliency of cyber-physical systems is different than conventional network-based cybersecurity.[2] While the intent of a robust cybersecurity program is good risk management, most traditional cybersecurity efforts are not holistic and tend to focus on protecting valuable assets.[15] Conversely, successful cyber-resiliency strategies involve understanding how mission-critical systems contribute to operational-level performance. For example, a mass transit system must safely transport people from one location to another without delay or cause for concern. While it might sound simple, this essential business operation requires several complex interactions among paying customers, supporting and enabling systems, the physical domain, and the information domain—all of which must be considered when designing for cyber resilience.

Table 2. Comparison of cyber-physical systems and traditional cyber attributes.

Comparative attributes	Cyber-physical systems	Traditional cyber attributes
Business advantage	Focused on real-time operations, ensuring the system is successful; considers what the business does to make a profit.	Focused primarily on protecting assets, mostly preventative with intense moments of reaction; considers what valuable business assets need to be protected.
Prioritization of the C-I-A triad (confidentiality, integrity, availability)	Focused on availability with assumed integrity and little regard for confidentiality.	Focused on retaining confidentiality of data along with integrity, and less priority on availability.
Scale of the complexity challenge	Complex interactions lead to poorly understood emergent behaviors.	Interactions might be complicated but are mostly linear, leading to well-understood behaviors.

Systems view: people, processes, and technology	These socio-technical systems require nearly constant inputs from users and sensors to monitor and control.	Mostly focused on technical security solutions and data security.

THE CYBER-RESILIENCY WORKFORCE

Having established a baseline understanding of cyber resiliency for cyber-physical systems, we now propose a basic job description to set expectations for the emerging cyber-resiliency workforce. Note that this description is meant for consideration by the broader cybersecurity community as the need for cyber-resiliency personnel is rapidly increasing across industry, academia, and government. Moreover, it is our hope that this discussion will help clarify the cyber-resiliency problem for cyber-physical systems.

> *Cyber-resiliency professionals provide holistic security solutions to ensure that mission-essential and safety-critical cyber-physical systems maintain effective functionality when operating despite facing adverse events.*

Workforce Responsibilities

Based on this job description and discussions with senior security personnel, we present several cyber-resiliency responsibilities. These responsibilities emphasize task execution and desired outcomes, and are organized such that job-specific duties can be understood more concretely. Additionally, these responsibilities can be used to inform education and training programs, career development paths, and roles within larger organizations.

1. Develop holistic, resiliency-informed system views that thoroughly account for the complexities and real-time operational constraints associated with operationally oriented cyber-physical systems.
2. Analyze the system's execution of essential business/mission operations in dynamic cyber-physical environments to include consequences from advanced cyberthreats, disruptions, disasters, and unpredictable emergent behaviors.
3. Define business/mission and system-level problem spaces accounting for cyber-related operational challenges and complex system-of-systems cyber dependencies.
4. Develop feasible resiliency strategies and objectives by considering current and future cyberthreat capabilities, criticality of the cyber-physical system's operation, and potential risks.
5. Perform security and resiliency requirements definition, engineering, and traceability tasks across the system's entire lifecycle.
6. Accomplish program management activities to ensure timely and integrated cybersecurity and resiliency solutions into program schedules, designs, and milestones.
7. Execute innovative engineering approaches toward the successful development, fielding, operation, and maintenance of secure and resilient cyber-physical systems.
8. Analyze potential solutions and their impact on personnel, processes, and technologies that reduce both technical and operational risk while meeting the system's performance expectations.
9. Perform tradeoff analysis of potential security and resiliency solutions for feasibility to include cost, performance, and schedule impacts.
10. Conduct testing activities that produce evidence of correct implementation of selected security and resiliency solutions.

Workforce Roles

In addition to the aforementioned responsibilities, we present several example cyber-resiliency work roles for consideration. While these work roles are largely dependent upon the experience level (novice, journeyman, or expert) and job-specific requirements, they are intended to illustrate how a cyber-resiliency professional might be required to communicate and interact with other personnel and organizations, including business/mission owners, managers, and other security specialists.

The cyber-resiliency workforce of the future will need to possess expertise in security, resiliency, relevant operational domains, and the system development lifecycle.

Sample expert-level work roles:

- Provide cyber-resiliency technical leadership to include considering competing designs, reviewing network architectures and documentation, and delivering presentations to a wide assortment of interested parties.
- Manage and/or interact with personnel (formally and informally) to provide technical guidance, mentoring, training, career development, and supervision.
- Initiate, build, and maintain relationships with key decision makers and stakeholders within and outside the organization.

Sample novice/journeyman-level work roles:

- Contribute as a team member to the development, analysis, and/or implementation of cyber-resiliency solutions.
- Inform team members of evolving security policies, standards, and approaches.
- Teach team members how to use existing and innovative cyber-resiliency tools, techniques, and procedures.

CYBER-RESILIENCY WORKFORCE EXPERTISE

Another way to examine the emerging need for resiliency-aware cybersecurity personnel is to consider the necessary areas of expertise for advanced cyber-physical systems, as shown in Figure 2. The cyber-resiliency workforce of the future will need to possess expertise in security and resiliency, relevant operational domains, and the system development lifecycle. The ideal cyber-resiliency professional would possess expertise in all three areas; however, we recognize that expertise requires years of experience to achieve (more than 10 years in some cases).[16] Thus, a more reasonable goal might be to require expertise in one area and familiarity in the other two. While it is difficult to determine which area is most important, modern decision makers often require security-focused personnel to make decisions in accordance with stated business goals and not the newest technological innovations.

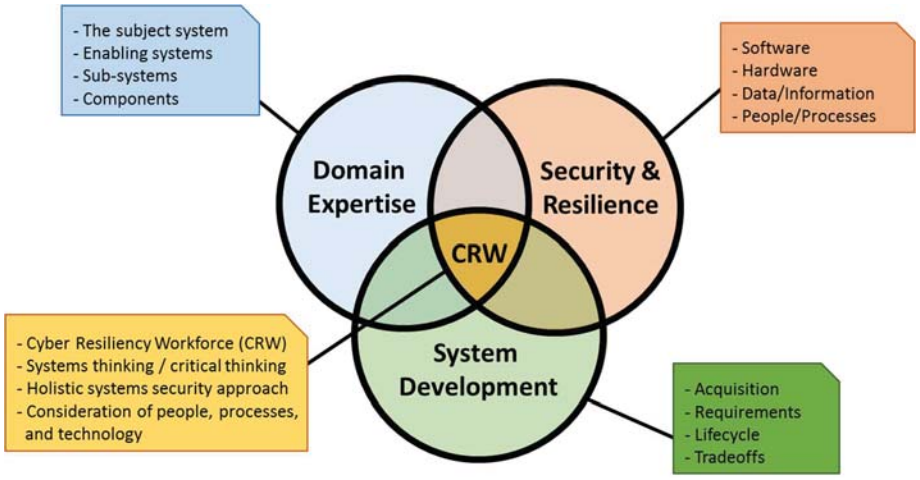


Figure 2. Multiple expertise areas are required for the cyber-resiliency workforce (CRW).

Building the Cyber-Resiliency Workforce

Assuming the cyber-resiliency workforce of the future will mature from the existing pool of cybersecurity personnel, the workforce must first learn essential resiliency concepts and principles. While there are hundreds (and potentially thousands) of sources available to understand and apply cybersecurity concepts, there are relatively few that discuss resiliency.[17] As a baseline, cyber-resiliency professionals should gain a working knowledge of system design, development, verification, and validation of software, hardware, and firmware.[9] Moreover, consideration for the appropriate level of technical and analytical rigor is important to ensure that cyber-resiliency solutions are feasible to key stakeholders.[9]

System development and program management experience is also critical because security and resiliency solutions are competing against other business/mission needs in resource-constrained environments. More so now than ever, the challenge of developing secure and resilient systems requires consideration across the system's entire lifecycle, from conceptual design phases to secure operation—including consideration for increasingly costly upgrades, modifications, recalls, and patches. Accordingly, proficiency with engineering processes, project milestones, design reviews, and decision criteria are necessary to design for resilient cyber-intensive systems.

Lastly, it is paramount for cyber-resiliency personnel to learn to communicate more effectively with domain-specific operators and managers. This is because domain expertise is critically important for gaining a fuller understanding of how to securely operate cyber-physical systems and meet key resiliency requirements.[18] For example, experience with aircraft operations (and/or maintenance) is essential for understanding how cyberthreats and incidents can negatively impact the system's ability to perform its intended mission.

Because this level of experience is rare, forward-looking cybersecurity professionals should actively seek out opportunities to gain familiarity and experience with the implementation of resiliency solutions in cyber-physical systems.[19]

DEVELOPING CYBER-RESILIENCY KNOWLEDGE, SKILLS, AND ABILITIES

To further understand the expertise required by cyber-resiliency professionals, we briefly discuss job-specific knowledge, skills, and abilities (KSAs). Although KSAs are often used in formal settings such as government employment, these three terms often cause confusion because they are used too informally. For example, while "knowledge" is fairly well understood as facts and

information, the differentiation between "skills" and "abilities" is less clear. Abilities tend to be inherent capabilities whereas skills are learned behaviors.

Research shows that expert performance is predominately recognized in the form of superior skills rather than knowledge gained or inherent ability. For example, the performance of an expert is most notable in their timeliness, consistency, and discernment.[17] More formally, we consider these skills as the acquired proficiency to perform job-related tasks such as programming secure software, executing test activities, and performing formal security analysis. Thus, those seeking to advance their cybersecurity careers should focus on improving their cyber-resiliency skills through new training opportunities and challenging work-related experiences. While the skills associated with resiliency are not well defined, ongoing work toward this goal is being performed by NIST through the National Initiative for Cybersecurity Education (NICE) framework in NIST SP 800-181, the NIST cyber-physical working group, and NIST SP 800-160 vol. 2, which is focused on resiliency. These timely efforts mean that individuals and companies have an excellent opportunity to learn and invest in the next generation of resiliency experts.

CONCLUSION

There is a rapidly growing opportunity for cybersecurity professionals to meet the needs of industry and government in the development of secure and resilient cyber-physical systems. This article uniquely details several key responsibilities, work roles, and expertise areas in order to prepare the cyber-resiliency workforce of the future. We provide direction to professionals seeking to advance their cybersecurity careers in resiliency and encourage motivated individuals to seek out additional opportunities to engage in challenging cyber-resilience experiences where possible.

DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the US Government.

REFERENCES

1. *Remarks by the President on Securing our Nation's Cyber Infrastructure*, The White House, 29 May 2009; https://obamawhitehouse.archives.gov/the-press-office/remarks-president-securing-our-nations-cyber-infrastructure.
2. Y. Liu et al., "Review on Cyber-Physical Systems," *IEEE/CAA J. Automatica Sinica*, vol. 4, no. 1, 2017, pp. 27–40.
3. R. Saracco, "Guess What Requires 150 Million Lines of Code," *EIT Digital*, blog, 13 January 2016; www.eitdigital.eu/news-events/blog/article/guess-what-requires-150-million-lines-of-code.
4. R.N. Charette, "This Car Runs on Code," *IEEE Spectrum*, 1 February 2009, pp. 3–7; http://spectrum.ieee.org/transportation/systems/this-car-runs-on-code.
5. A. Greenberg, "Hackers Remotely Kill a Jeep on the Highway—With Me in It," *Wired*, blog, 21 July 2015; www.wired.com/2015/07/hackers-remotely-kill-jeep-highway.
6. E. Perez, "FBI: Hacker Claimed to Have Taken over Flight's Engine Controls," *CNN*, 18 May 2015; www.cnn.com/2015/05/17/us/fbi-hacker-flight-computer-systems.
7. S. Checkoway et al., "Comprehensive Experimental Analyses of Automotive Attack Surfaces," *Proc. 20th USENIX Conf. Security* (SEC), 2011; https://dl.acm.org/citation.cfm?id=2028073.
8. M. Maybury, "Toward the Assured Cyberspace Advantage: Air Force Cyber Vision 2025," *IEEE Security & Privacy*, vol. 13, no. 1, 2015, pp. 49–56.
9. R.S. Ross, M. McEvilley, and J.C. Oren, *Systems Security Engineering: Considerations for a Multidisciplinary Approach in the Engineering of Trustworthy Secure Systems*, government report SP 800-160, 2016;

www.nist.gov/publications/systems-security-engineering-considerations-multidisciplinary-approach-engineering.
10. D. Bodeau et al., *Cyber Resiliency Engineering Framework*, technical report MTR110237, MITRE Corporation, 2012; www.mitre.org/publications/technical-papers/cyber-resiliency-engineering-framework.
11. S. Hosseini, K. Barker, and J.E. Ramirez-Marquez, "A Review of Definitions and Measures of System Resilience," *Reliability Engineering & System Safety*, vol. 145, 2016, pp. 47–61.
12. H. Goldman, R. McQuaid, and J. Picciotto, "Cyber Resilience for Mission Assurance," *IEEE Intl. Conf. Technologies for Homeland Security* (HST), 2011; doi.org/10.1109/THS.2011.6107877.
13. T. Benzel, "A Strategic Plan for Cybersecurity Research and Development," *IEEE Security & Privacy*, vol. 13, no. 4, 2015, pp. 3–5.
14. *INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*, Wiley, 2015.
15. L. Hoffman, D.H. Burley, and C. Toregas, "Holistically Building the Cybersecurity Workforce," *IEEE Security & Privacy*, vol. 10, no. 2, 2012, pp. 33–39.
16. M.J. Assante and D.H. Tobey, "Enhancing the Cybersecurity Workforce," *IT Professional*, vol. 13, no. 1, 2011, pp. 12–15.
17. *Cyber Resiliency Resource List*, blog, MITRE Corporation, 4 May 2016; www2.mitre.org/public/sr/Cyber-Resiliency-Resources-16-1467.pdf.
18. *Report on Securing and Growing the Digital Economy*, report, Commission on Enhancing National Cybersecurity, 1 December 2016; www.nist.gov/sites/default/files/documents/2016/12/02/cybersecurity-commission-report-final-post.pdf.
19. S. Musman, "Assessing Prescriptive Improvements to a System's Cyber Security and Resilience," *2016 Ann. IEEE Systems Conf.* (SysCon), 2016; doi.org/10.1109/SYSCON.2016.7490660.

ABOUT THE AUTHORS

Logan O. Mailloux serves as a commissioned officer in the United States Air Force and works as an assistant professor at the Air Force Institute of Technology (AFIT) at Wright-Patterson Air Force Base. His research interests include system security engineering, complex information communication and technology implementations, and quantum key distribution systems. Mailloux is a Certified Information Systems Security Professional and a Certified Systems Engineering Professional. He received a PhD in systems engineering from AFIT and is a member of Tau Beta Pi, Eta Kappa Nu, the International Council on Systems Engineering, and IEEE. Contact him at logan.mailloux@us.af.mil.

Michael R. Grimaila is a professor of systems engineering, head of the Department of Systems Engineering and Management, and a member of the Center for Cyberspace Research at AFIT. His research interests include computer engineering, mission assurance, quantum communications and cryptography, data analytics, network management and security, and systems engineering. Grimaila is a Certified Information Security Manager and Certified Information Systems Security Professional. He is a member of Tau Beta Pi, Eta Kappa Nu, and ACM; a senior member of IEEE; and a fellow of the Information Systems Security Association. Contact him at michael.grimaila@afit.edu.

The New Threats of Information Hiding

The Road Ahead

Krzysztof Cabaj
Warsaw University of Technology

Luca Caviglione
National Research Council of Italy

Wojciech Mazurczyk
Warsaw University of Technology

Steffen Wendzel
Worms University of Applied Sciences

Alan Woodward
University of Surrey

Sebastian Zander
Murdoch University

A recent trend involves exploiting various information-hiding techniques to empower malware—for example, to bypass mobile device security frameworks or to exfiltrate sensitive data. The authors provide an overview of information-hiding techniques that can be utilized by malware. They showcase existing and emerging threats that use different types of data-hiding mechanisms (not just those adopting classical covert channels), with the goal of monitoring these threats and proposing efficient countermeasures.

The use of information-hiding techniques, often referred to as steganography, to commit cyberattacks or crimes has received relatively little attention in the academic literature or the media. When mentioned, steganography is typically discussed in the context of covert communication between extremist individuals or groups.[1] Even then, some argue that there is little or no evidence that steganography is in use. While large-scale surveys found no conclusive traces of the use of data hiding, some researchers warn against concluding that it is not in use.[2]

Recently, there have been signs that things are starting to change. Reports from McAfee[3] and Kaspersky[4] recognized the role that information hiding plays in current malicious software and that it is highly likely to gain additional importance in the future. Furthermore, because of the sensitivity of the subject, organizations are often reluctant to report the detected use of steganography to the public.[5]

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

Historically, cryptography has been a more widely discussed topic than steganography, especially in law enforcement. In the past, the mere existence of encrypted communications and data would have raised suspicions, but it is a frequent scenario today. For example, malware using encrypted communications for command and control (C&C) purposes might previously have stood out from regular network traffic, but now it is effectively hidden within the "background noise" of routinely encrypted data exchanged in the network. Nevertheless, encrypted communications can be detected relatively easily, and ancillary techniques—such as traffic analysis or metadata recovery—allow for at least some intelligence to be derived from encrypted data and communications. The recovered metadata (such as who is communicating with whom, when, and for how long) can be as or even more important than knowing the actual content.

Currently, encryption is receiving greater attention from security professionals, law enforcement, and security and intelligence agencies. For example, recent advancements in understanding how malicious software encrypts its own communications could help identify and block C&C communications of botnets.[6] Unfortunately, criminals or extremists are well aware of the increased focus on encryption and are looking for other ways to make malicious software stay under the radar, especially in the context of stealing data (where triggering some form of defense must be avoided). In this vein, the most important and recent trend is to equip malware with information-hiding capabilities, or techniques that hide communications.[7]

This article provides an overview of information-hiding techniques that can be utilized by malware. By using real-world examples, this article showcases existing and emerging threats using different types of data-hiding mechanisms (not just those adopting classical covert channels). The research presented here was performed within the Criminal Use of Information Hiding (CUING) initiative (http://cuing.org), which was formed with the cooperation of the Europol European Cyber Crime Centre (EC3) to gather experts from different backgrounds with the aim of monitoring information-hiding-capable threats and proposing efficient countermeasures.

COVERT CHANNELS AND DATA HIDING

Cyberattacks are commonly divided into five phases:[8] reconnaissance (gathering information), scanning the target, gaining access to the target, maintaining access, and covering the tracks. Information-hiding techniques are mostly applied in phases 2 to 4, on which we focus here. Figure 1 shows the classification of information-hiding techniques and how they are used by malware in different attack phases.

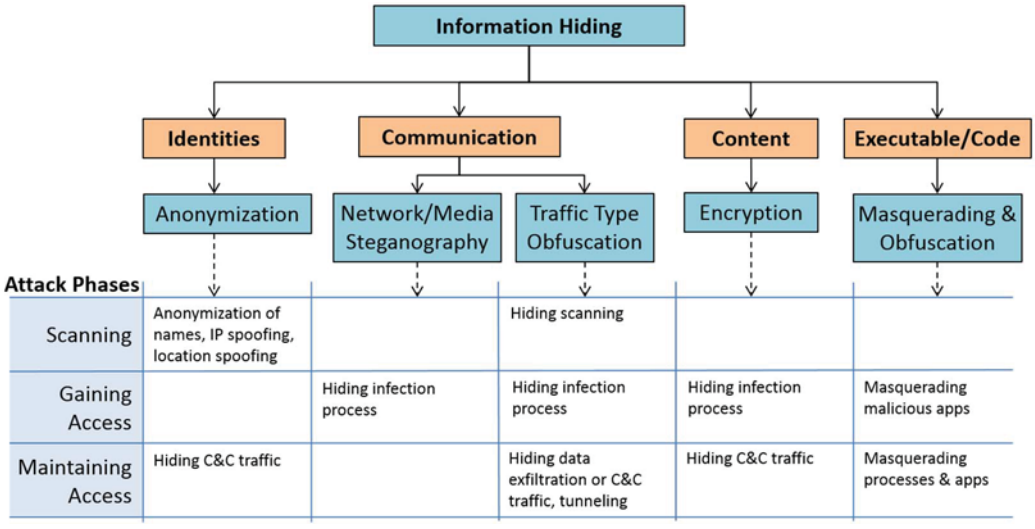


Figure 1. Classification of information-hiding techniques. C&C: command and control.

As depicted in Figure 1, information hiding is a very broad term. It encompasses different subdisciplines (or domains), which can be used by an attacker during different attack stages depending on what is subjected to hiding, including the following.

- *Identities*. The identities of communicating parties are hidden by anonymization techniques.
- *Communication.* The fact that a communication is taking place is hidden by steganography techniques. The characteristics of a network conversation (for example, a packet flow) can be concealed using traffic-type obfuscation methods.
- *Content*. Hiding the content of data but not the transmission or presence of the data itself is achieved by applying cryptographic algorithms.
- *Code*. The structure of (executable) code is hidden by (binary) code obfuscation and masquerading techniques.

First, let us discuss the most important (from our perspective) data-hiding methods—those that conceal the fact that a communication is taking place. Typically, this type of information hiding is realized using some form of steganography.

Historically, the earliest computer steganographic methods were focused on different media types, especially digital images. For example, several algorithms hide information within the least-significant bits (LSBs) of color definitions of pixels within an image, as the human eye cannot spot such alterations. A similar approach has been used for audio and video. The natural evolution is to hide data in network transmissions, such as in inter-arrival times of packets or in unused fields of protocol headers. Network traffic provides the advantage of a continuous data flow, which a digital media file of constant size cannot provide. When secret data is hidden in network traffic, the secret communication channel is referred to as a network covert channel.

In essence, network covert channels enable secret malware communications over any type of computer network, be it a local area network or the Internet. Compared to encryption, which only ensures the confidentiality of what a malware communicates, covert channels can help keep the communication secret and to retain access to a hacked system. Moreover, control protocols can be used on top of covert channels, representing a form of C&C channel. Such control protocols allow attackers to upload a newer version of a malware binary, to select a different encryption or covert signing scheme, to switch from one steganographic method to another, or to apply dynamic overlay routing to bypass firewalls.[9] Malware can also apply network covert channels to conceal the exfiltration of organizational data over the network and to bypass firewalls by hiding data in transmissions that are not affected by its filtering policy. These goals especially affect phases 2 and 3 (gaining/maintaining access). Note that when referring to malware trying to communicate covertly or abuse some network service, the hacking community often uses the term "tunneling." However, this is not accurate because tunneling hides traffic as a byproduct, and actually refers to the encapsulation of network data of the same or higher layer—for example, IPv4 as payload in an IPv6 packet.

While steganography aims to hide data inside digital objects, two other classes of methods obfuscate information in code (code obfuscation) or network traffic (traffic-type obfuscation). Obfuscation is different from steganography—the latter tries to communicate secret data in a nonnoticeable manner while the former is directly visible to an analyst. Despite their different strategies, both domains share the goal of hiding data. The goals of traffic type and code obfuscation affect phase 1 (scanning), but mainly affect phases 2 and 3 (gaining/maintaining access).

Anonymization provides a means of communication without revealing private attributes of the communicating peers, such as their names, IP addresses, or geographical locations. In contrast to steganography, anonymization relies on different techniques—such as spoofing the IP address of a sender or cryptographic algorithms—to fake or hide sensitive data that can be used to deduce information about the parties involved in a communication. Note that, as shown in Figure 1, cryptographic methods can be used to encrypt any kind of secret data, not just data that reveals identities. Thus, the application of cryptography is not limited to anonymity techniques. Anonymity techniques can be utilized during phases 1 (scanning) and 3 (maintaining access), while encryption (despite its use for anonymity purposes) affects phases 2 and 3 (gaining/maintaining access).

INFORMATION-HIDING MALWARE IN THE WILD

Here we present several examples of information-hiding malware observed in the wild. Because of space constraints, we focus only on the most representative threats observed from 2011 to 2017.

Originally, information-hiding techniques were implemented only in advanced persistent threats (APTs) like Duqu, Regin, or Hammertoss—the most sophisticated types of malware created with the support of nationwide sponsors. However, information-hiding techniques are slowly becoming the de facto standard for "ordinary" malware. For example, various types of popular threats like ransomware (TeslaCrypt, Cerber, and SyncCrypt) or exploit kits (Stegano/Astrum, DNSChanger, and Sundown) use some form of information hiding. Examples of existing information-hiding malware are summarized in Table 1.

Table 1. Main examples of existing information-hiding malware.

Malware/exploit kit	Information-hiding method	Purpose
Vawtrak/Neverquest	Modification of the least-significant bits (LSBs) of favicons	Hiding URL to download a configuration file
Zbot	Appending data at the end of a JPG file	Hiding configuration data
Lurk/Stegoloader	Modification of the LSBs of BMP/PNG files	Hiding encrypted URL for downloading additional malware components
AdGholas	Data hiding in images, text, and HTML code	Hiding encrypted malicious JavaScript code
Android/Twitoor.A	Impersonating a pornography player or an MMS app	Tricking users into installing malicious apps and spreading infection
Fakem RAT	Mimicking MSN and Yahoo Messenger or HTTP conversation traffic	Hiding command and control (C&C) traffic
Carbanak/Anunak	Abusing Google cloud-based services	Hiding C&C traffic
SpyNote Trojan	Impersonating Netflix app	Tricking users into installing malicious app to gain access to confidential data
TeslaCrypt	Data hiding in HTML comments tag of the HTTP 404 error message page	Embedding C&C commands
Cerber	Image steganography	Embedding malicious executable
SyncCrypt	Image steganography	Embedding core components of ransomware

Stegano/Astrum	Modifying the color space of the used PNG image	Hiding malicious code within banner ads
DNSChanger	Modification of the LSBs of PNG files	Hiding malware AES encryption key
Sundown	Hiding data in white PNG files	Exfiltrating user data and hiding exploit code delivered to victims

Malware Using Modifications to Digital Media Files

Currently, one of the most common ways to hide data is to use digital media files as the secret carrier. The most common technique exploits digital images to do one of the following: conceal malware settings or a configuration file, provide the malware with a URL from which additional components can be downloaded, or directly store the whole malicious code. The most notable example took place in 2015 when Vawtrak/Neverquest malware started utilizing steganography to hide settings in favicons (innocent-looking pictures widely available on websites). The malware extracts the LSBs from each image's pixel to reconstruct a previously embedded URL for downloading its configuration file. A similar approach has been used by Zbot malware, which downloaded an innocent-looking JPEG image on the infected system containing its configuration data appended at the end of the image. Lurk and Stegoloader used the LSB of a digital image (BMP and PNG, respectively) to retrieve an encrypted URL for downloading additional software components.

More recently, we observed the use of information-hiding techniques for malvertising (malicious advertising) attacks as evidenced by the AdGholas malware. AdGholas avoids detection by using steganography for hiding encrypted JavaScript code in images, text, and HTML code. At the end of 2016, large-scale attacks related to the online e-commerce platform Magento revealed the use of image steganography to conceal payment card details. Once the platform was infected, the malware collected payment details and hid them inside images of real products available on the infected e-commerce site. By downloading such modified images, the attacker could easily exfiltrate the stolen data.

Malware Posing as Other Legitimate Applications or Mimicking Their Traffic Behavior

Some malware relies on the mimicry of legitimate programs and/or their communications. A paradigmatic example is a variant of Android/Twitoor.A—malware that spreads by SMS or malicious URLs. The malware impersonates a pornography player or an MMS application but without the correct functionality, eventually tricking the user to install the application and spread the infection. Another application, Irongate, is the first notable example designed to operate in industrial control systems scenarios. One of the most important features is its ability to record several seconds of ordinary, legitimate traffic from a programmable logic controller and then use it as a smokescreen (in other words, the malicious commands are masked using legitimate ones) when sending intentionally modified data back. Such an operation allows the attacker to alter a controlled process without raising any security alerts. Another example is Fakem RAT, which made its C&C traffic look like MSN and Yahoo Messenger or HTTP conversations.

At the beginning of 2017, Carbanak/Anunak demonstrated its ability to abuse Google cloud-based services to set up a covert channel for C&C purposes. In this case, a unique Google Sheets spreadsheet was dynamically created to manage each infected victim. The use of a Google service granted attackers the ability to stay under the radar because such third-party services are typically not blocked in the enterprise network and are considered safe. Another example includes a new version of SpyNote Trojan, which was disguised as a legitimate Netflix application. Once installed, it allowed the attacker to execute different actions, such as copy a user's files, view a user's contacts, and eavesdrop on a user's communication.

A technique called domain fronting is gaining a lot of attention, especially among APT-related groups. Put briefly, it is used to mask the true destination of a connection by mimicking legitimate traffic to an innocent destination. A successful implementation exploits HTTPS traffic to communicate with an infected host, making the traffic look like a Google search. Instead, the traffic is produced by a connection exchanging data with the attacker.

Information Hiding in Ransomware

The first instances of using ransomware to hide information were discovered at the beginning of 2016 when TeslaCrypt was spread using the Neutrino exploit kit. Neutrino initially redirects users to a malicious landing page crafted for discovering the victim's vulnerabilities to deliver the most appropriate exploit. If the vulnerability is successfully exploited, a downloader is executed. To gather data, it contacts a server, which responds with an HTTP 404 error page that embeds C&C commands in the HTML comments tag.

In mid-2016, Cerber was identified as one of the macro-type malware-delivered ransomware across a variety of cloud-based file-sharing applications. To spread the infection, Cerber uses a decoy document which, when opened, loads a malicious macro-code that downloads a JPEG file to the targeted machine. Inside this benign-looking image is the steganographically embedded malicious executable.

In August 2017, a similar technique was discovered with the SyncCrypt ransomware. Infected emails contain Windows Script File (WSF) attachments posing as court orders. If opened, malicious code downloads a digital image containing the core components of SyncCrypt.

Information Hiding in Exploit Kits

Information-hiding methods have become so popular among cybercriminals that they are now incorporated within exploit kits to allow developers with little or no programming skills to create, customize, and distribute malware. The first example of this is the Stegano/Astrum exploit kit, which was used in 2016 as part of a huge malvertising campaign. Malicious code is embedded within banner ads by modifying the color space of the used PNG image (the alpha channel). Then, the victim's browser parses an injected JavaScript code, extracting the malicious code and redirecting users to the exploit kit landing page. The infection is performed on the landing page, typically by using several Flash vulnerabilities.

DNSChanger, another type of malvertising exploit kit identified in 2016, hides an AES encryption key within an innocent-looking ad to decrypt the network traffic generated by the exploit kit. The scope of DNSChanger is to launch brute-force attacks against the network routers to take control of the victim's network and inject ads in all exchanged traffic.

While Stegano/Astrum and DNSChanger are niche products, the Sundown exploit kit is one of the major players in the exploit kit market. Sundown uses steganography in two ways: to covertly exfiltrate information stolen from the infected system in PNG files (which are uploaded to an Imgur album where cybercriminals can access them undisturbed—see the CryLocker ransomware campaign as an example) and to hide the exploited code delivered to the victims.

THE ROAD AHEAD

We have experienced a massive growth in cybercrime in recent years, and this trend is likely to continue because it can be so lucrative.[3] We see the following main developments in cybercrime: increased stealth, commoditization of malware, and exploitation of Internet of Things (IoT) devices. Cybercriminals will place more emphasis on making it harder to detect and trace back malware to its origin, which will be a main driver for the increased use of information hiding.

Because a main goal of malware developers is to always remain one step ahead, they will continually try to improve their information-hiding techniques. One avenue is to utilize better digital media steganography algorithms. Improved algorithms, which are harder to detect and eliminate,

are already available and known among academics (for example, F5[10] and HUGO[11]). Another strategy is to hide information in new services or protocols such as Skype, [12] BitTorrent, and Stream Control Transmission Protocol (SCTP).[13] When these are targeted, the result is a "needle in a haystack" problem when it comes to detecting covert communications among a large number of similar connections.

Another future direction is to exploit the ongoing IPv4 to IPv6 transition. Malware can take advantage of misconfigured nodes or hosts with IPv4-only stacks that are unable to process IPv6 malicious traffic. Malware also increasingly exploits the diffusion of HTTPS by hiding in HTTPS or Transport Layer Security (TLS) traffic, which cannot be easily inspected (researchers claim that one-third of malware already uses HTTPS).[14]

Botnets will remain an important tool for cybercriminals for various purposes, such as managing distributed denial of service (DDoS) attacks or sending spam emails. Because bots can be relatively easily identified by observing the C&C traffic, masquerading this traffic is very important. While most existing approaches are simple (for example, C&C protocols hide in HTTP, IRC, or DNS), researchers recently demonstrated how to completely transform a C&C protocol to mimic another innocuous protocol.[15] Future botnets might utilize overlay networks that use only steganographic methods to communicate (stego-botnets).[16]

The DNS protocol is a natural choice to hide C&C traffic or for data exfiltration as it cannot be blocked. Developing stealthier covert channels on top of DNS and developing the countermeasures to detect these channels is an ongoing arms race[17] that could become even more interesting once the Domain Name System Security Extensions (DNSSEC) are more widely deployed.

Future attacks will target the ever-increasing number of IoT devices, such as networked sensors, CCTV cameras, smart TVs and DVRs, smart home and building appliances, and industrial control systems. In many cases IoT devices are soft targets, as their processing capabilities limit the implemented security mechanisms and the low cost of many of these devices means that security is often an afterthought for manufacturers. Attacks on IoT devices also allow user profiling and maliciously interfering with the physical world. Moreover, malware can utilize the IoT to hide secret data.[18] For instance, an attacker can secretly store data in unused registers of IoT devices or by slightly modifying actuator states.[19]

Currently deployed steganography methods are often simple, mainly because current protection solutions (such as intrusion detection systems) hardly detect any form of steganography in practice. Thus, malware developers are not forced to apply more sophisticated steganography. Nevertheless, recent threats often merge simple covert channel techniques with memory-resident or fileless implementations to make them stealthier and able to cover their tracks on the infected host—for example, in the filesystem.

However, data loss prevention (DLP) solutions increasingly aim to detect steganographic transmissions. This will force malware authors to improve the covertness of their data-leakage techniques. That said, cybercriminals will increasingly choose off-the-shelf malware rather than develop custom malware, which would require more financial investment. Once more advanced steganography finds its way into off-the-shelf malware products, such as exploit kits, it will become widely used at relatively little extra cost to the cybercriminals.

When the volume of more sophisticated malware increases, malware de-obfuscation and steganography analysis must be done in a more systematic and efficient way. Frameworks for distributed and automated malware analysis like the Malware Analysis and Storage System (MASS) could be a suitable approach for handling large volumes of malware samples retrieved from honeynets.[20]

CONCLUSIONS AND FUTURE WORK

Modern malware has become so efficient that it can remain covert for a long time. Even if steganographic techniques are not the main reason for this efficiency, the ability to create and exploit covert channels for C&C and exfiltration purposes surely plays a role (for example, Regin went undiscovered from 2008 to 2014). This fact is exacerbated by a worrying lack of techniques for detecting information-hiding threats, especially regarding the IoT and automation. A possible

cause is the poor generalizability of the process of detecting information hiding. Many detection techniques are tightly coupled with specific hiding methods, the cover they use, the scenario in which they are used, and the technology on which they depend.

Because creating new hiding methods by applying known techniques to new protocols, scenarios, and technology is relatively easy, countermeasures are always at least one step behind. Therefore, industry and academia should focus on the development of new and general tools or add-ons for the most common network security solutions. One idea is the use of new and more general indicators, such as patterns used by different hiding techniques or energy consumption.

Information hiding increases the complexity of addressing cybersecurity. Organized initiatives like CUIng can be the incubator where a long-term cure for information-hiding malware is developed, as modern cyberthreats require a multidisciplinary approach with the collaboration of many experts from industry, academia, and law-enforcement agencies.

REFERENCES

1. D. Anderson, *A Question of Trust: Report of the Investigatory Powers Review*, government report, Independent Reviewer of Terrorism, June 2015; https://terrorismlegislationreviewer.independent.gov.uk/wp-content/uploads/2015/06/IPR-Report-Print-Version.pdf.
2. N. Provos and P. Honeyman, "Hide and Seek: An Introduction to Steganography," *IEEE Security & Privacy*, vol. 1, no. 3, May 2003, pp. 32–44.
3. *McAfee Labs Threats Report*, report, McAfee, June 2017; http://mcafee.ly/2sXowrq.
4. A. Shulmin and E. Krylova, "Steganography in Contemporary Cyberattacks," *SecureList*, blog, 3 August 2017; https://securelist.com/steganography-in-contemporary-cyberattacks/79276/.
5. J. Millen, "20 Years of Covert Channel Modeling and Analysis," *Proc. 1999 IEEE Symp. Security and Privacy*, 1999; doi.org/10.1109/SECPRI.1999.766906.
6. L. De Carli et al., "Botnet Protocol Interference in the Presence of Encrypted Traffic," *IEEE Conf. Computer Communications* (INFOCOM), 2017; doi.org/10.1109/INFOCOM.2017.8057064.
7. W. Mazurczyk and L. Caviglione, "Information Hiding as a Challenge for Malware Detection," *IEEE Security & Privacy*, vol. 13, no. 2, 2015, pp. 89–93.
8. *Ethical Hacking and Countermeasures: Attack Phases*, EC-Council, Course Technology, 2009.
9. W. Mazurczyk et al., *Information Hiding in Communication Networks: Fundamentals, Mechanisms, Applications, and Countermeasures*, Wiley-IEEE Press, 2016.
10. A. Magnúsardóttir, "Malware is Heavily Moving to HTTPS," *Cyren Security Blog*, blog, 7 June 2017; https://blog.cyren.com/articles/over-one-third-of-malware-uses-https.
11. W. Mazurczyk and S. Wendzel, "Information Hiding: Challenges for Forensic Experts," *Communications of the ACM*, vol. 61, no. 1, 2018, pp. 86–94.
12. S. Wendzel, W. Mazurczyk, and G. Haas, "Steganography for Cyber-physical Systems," *J. Cyber Security and Mobility*, vol. 6, no. 2, 2017, pp. 105–126.
13. A. Westfeld, "F5--A Steganographic Algorithm," *Information Hiding*, Springer, 2001.
14. T. Pevný, T. Filler, and P. Bas, "Using High-Dimensional Image Models to Perform Highly Undetectable Steganography," *Information Hiding*, Springer, 2010.
15. W. Mazurczyk, "VoIP Steganography and its Detection—A Survey," *ACM Computer Surveys*, vol. 46, no. 2, 2013; doi.org/10.1145/2543581.2543587.
16. W. Fraczek, W. Mazurczyk, and K. Szczypiorski, "Hiding Information in Stream Control Transmission Protocol," *Computer Communications Journal*, vol. 35, no. 2, 2012, pp. 159–169.
17. P. Backs, S. Wendzel, and J. Keller, "Dynamic Routing in Covert Channel Overlays Based on Control Protocols," *2012 Int'l Conf. Internet Technology and Secured Transactions*, 2012, pp. 32–39.
18. X. Zhong et al., "Stealthy Malware Traffic - Not as Innocent as it Looks," *10th International Conf. Malicious and Unwanted Software* (MALWARE), 2015, pp. 110–116.

19. S. Sheridan and A. Keane, "Improving the Stealthiness of DNS-Based Covert Communication," *16th European Conf. Cyber Warfare and Security* (ECCWS), 2017; https://arrow.dit.ie/nsdcon/3.
20. F. Rump, T. Behner, and R. Ernst, "Distributed and Collaborative Malware Analysis with MASS," *IEEE 42nd Conf. Local Computer Networks* (LCN), 2017, pp. 191–194.

ABOUT THE AUTHORS

Krzysztof Cabaj is an assistant professor at Warsaw University of Technology (WUT) and an instructor of Cisco Academy courses at the International Telecommunication Union Internet Training Centre (ITU-ITC). His research interests include network security, honeypots, and data-mining techniques. Cabaj received a PhD in computer science from WUT. He is the author or co-author of more than 40 publications in the field of information security. Contact him at kcabaj@ii.pw.edu.pl.

Luca Caviglione is a researcher with the Institute of Intelligent Systems for Automation at the National Research Council of Italy. His research interests include P2P systems, wireless communications, cloud architectures, and network security. Caviglione received a PhD in electronics and computer engineering from the University of Genoa. Contact him at luca.caviglione@cnr.it.

Wojciech Mazurczyk is an associate professor at the Institute of Telecommunications at WUT. His research interests include network security, information hiding, and network forensics. Mazurczyk received PhD and DSc degrees in telecommunications from WUT. He is also an associate technical editor for *IEEE Communications Magazine*. Contact him at wmazurczyk@tele.pw.edu.pl.

Steffen Wendzel is a professor of information security and computer networks at Worms University of Applied Sciences. His research interests include information hiding, network security, and security in the Internet of Things—he has written five books on these topics. Wendzel received a PhD in computer science from the University of Hagen. Contact him at wendzel@hs-worms.de.

Alan Woodward is a visiting professor in the Centre for Cyber Security at the University of Surrey. He worked for the UK government for many years and continues to provide advice to government organizations through advisory roles. Woodward's research interests include cybersecurity, covert communications, forensic computing, and image/signal processing. He received a BSc in physics from the University of Southampton. Woodward is a Fellow and chartered member of the British Computer Society, Institute of Physics, and the Royal Statistical Society. Contact him at alan.woodward@surrey.ac.uk.

Sebastian Zander is a lecturer at Murdoch University. His research interests include network and computer security, covert channels, the IPv4 to IPv6 transition, transport protocols, and network measurement. Zander received a PhD in telecommunications engineering from Swinburne University of Technology. He is a member of the Australian Computer Society (ACS). Contact him at s.zander@murdoch.edu.au.

Internet of Things Forensics

The Need, Process Models, and Open Issues

Maxim Chernyshev
Edith Cowan University

Sherali Zeadally
University of Kentucky

Zubair Baig
CSIRO

Andrew Woodward
Edith Cowan University

The Internet of Things (IoT) brings a set of unique and complex challenges to the field of digital forensics. To take advantage of the volume and variety of data captured by and stored in ubiquitous IoT services, forensic investigators need to draw upon evidence-acquisition methods and techniques from all areas of digital forensics and possibly create new IoT-specific investigation processes. Although a number of conceptual process models have been developed to address the unique characteristics of the IoT, many challenges remain unresolved.

Recent advances in hardware, software, and communication technologies have accelerated the deployment of a wide range of Internet-enabled devices, resulting in the Internet of Things (IoT). Industry predictions indicate that the number of connected devices has already surpassed the population of the planet, with no foreseeable slowdown in growth.[1] The evolution and growth of the IoT has driven the convergence of several technological paradigms comprising wireless sensor networks (WSNs), mobile and cloud computing, and the Internet as the overarching ubiquitous connectivity enablers.

Environmental data–collection systems using sensors and physical world interactions via actuators enable valuable and convenient consumer and industry-focused applications and services. The current landscape of the IoT represents a ubiquitous, constantly evolving, pervasive, and highly heterogeneous network of interconnected devices with diverse physical properties and computational capabilities that are being deployed on a large scale for various applications such as healthcare, manufacturing, construction, automotive, retail, and engineering.

Unfortunately, security considerations are not always given sufficient priority during IoT system design and development. Inherent vulnerabilities in communication protocols and software stacks leave many devices susceptible to threats.[2] Cybercriminals exploit these vulnerabilities and continue to launch highly disruptive and large-scale attacks with increasing levels of sophistication. A case in point was the 2016 denial of service (DoS) cyberattack against Dyn domain

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

name servers.[3] Having the practical ability to investigate IoT-related cybercrimes will enable the successful and timely prosecution of those responsible, which is paramount to curbing the growth of adversarial threats.

The science of digital forensics focuses on supporting investigations involving digital devices, including those in the IoT ecosystem. Digital forensics relies on digital evidence, scientifically derived and proven evidence-acquisition methods, and validated tools used by qualified forensic experts. The main objective of digital forensics is to facilitate acquisition and analysis of forensically sound digital evidence that can be presented and admitted in a court of law. The emergence of the IoT has brought many challenges to digital forensics, especially by requiring current methods and techniques to be applied to a highly diverse and ever-changing digital environment.

The science of digital forensics focuses on supporting investigations involving digital devices, including those in the IoT ecosystem.

In this work, we present a succinct review of the state of the art of conceptual digital forensic models that can be applied to the IoT environment. We also discuss open issues that exist in these conceptual digital forensic techniques when they are applied to IoT devices.

THE NEED FOR IOT FORENSICS

The emergence of the IoT is perceived as a potential enabler for novelty in the digital investigation process. For instance, the data collected and shared by ubiquitous sensors present an abundance of potential digital evidence by virtue of their numbers, variety, and coverage in many application areas. The digital artifacts found in the IoT ecosystem can be used to support or refute investigation hypotheses and subsequently any claims made by parties involved in an investigation. In fact, we have already observed civil and criminal investigations that made use of data from consumer wearable devices in personal injury and murder cases.[4]

However, the underlying complexity involved in extracting data from the IoT infrastructure and its devices can hinder the investigator's ability to produce forensically sound and admissible evidence.[5] This complexity stems from a number of challenges outlined by R.C. Hegarty and colleagues:[6]

- uncertainty around where the data came from, where and how it is stored, and the data attributes that are stored;
- difficulty in securing the chain of custody due to increasing data volatility and complex data transit routes among the IoT architecture layers;
- inapplicability of traditional digital forensics extraction techniques to aggregated data stored in the cloud; and
- diverse and proprietary data storage and exchange formats featuring reduced granularity due to capacity constraints used by IoT services.

To illustrate the unique characteristics of IoT-based digital investigations, we discuss some of these challenges in this article. In the context of the IoT, an investigator will most likely need to examine a diverse range of potential evidence sources. This need often presents the requirement to select and combine multiple digital forensic methods and techniques, which increases the overall complexity of the investigation procedure as well as the baseline training, skill, and expertise requirements.

For example, consider a contemporary smart home. The Law Enforcement Cyber Centre's IoT infographic[7] includes 17 potential sources of digital evidence, including smart appliances, connected vehicles, personal assistants, personal health and medical devices, digital photo frames, smart meters, and home automation systems. These IoT sources run on a heterogeneous set of technologies that include a combination of multi-protocol wired and wireless communications, speakers, cameras, microphones, remote and local storage, ambient sensors, voice recognition, and location tracking. Extraction of evidence from all such devices requires currency in expertise across multiple digital forensic branches, such as computer, mobile, and embedded forensics for

working with local storage; network forensics for accessing and analyzing the communications medium; and cloud forensics for analyzing the remote storage. This issue becomes increasingly significant in the context of large-scale IoT environments such as industrial deployments and smart cities. These environments present a considerably larger number of potential individual evidence sources and introduce additional complexity due to significant technological diversity.

Furthermore, traditional digital forensic methods and techniques such as carving, which is used to search for specific content in an extracted filesystem image, might not apply to IoT devices such as lightweight sensors that rely on flash memory with no built-in filesystem storage capability.[8] Even if a flash memory image of a device such as a wireless sensor could be acquired by a known forensic data acquisition tool, it is unlikely that this tool or other tools concerned with image parsing would be able to interpret the underlying format correctly. Subsequently, the ability to produce human-readable evidence from IoT devices can be severely limited due to lack of consistency in format and protocol support. Although specialized tools and techniques can be developed to extract and interpret the contents of specific system on chip (SoC) circuit boards (for instance, to extract network topology–based evidence such as routing information[9]), derivation and validation of SoC-specific techniques can be a slow process, which can prove to be unsustainable in practice.

The forensic tools taxonomy provided by NIST does not clearly identify the tools that can be useful in an IoT-based investigation. The taxonomy lists only a handful of tools (such as iVe and XRY Complete) that target embedded devices that are widely used in the IoT sensors landscape outside of the consumer segment.[10] Unlike consumer-grade connected devices such as smart appliances (like smart televisions and refrigerators), embedded IoT sensors and actuators are usually based on low-power constrained chips that are based on specialized energy-efficient routing and application layer protocols such as the Routing Protocol for Low-Power and Lossy Networks (RPL) and the Constrained Application Protocol (CoAP) for connectivity and data transfers.[11] The limited capabilities of resource-constrained IoT devices naturally result in higher data volatility. As a result, potential evidence might not be present at all on the device or might only reside on the device for a very short period of time before being overwritten by more recent data.

Finally, the complexity of digital forensic investigations increases with the number of potential evidence sources. Consider a compromised IoT solution comprising multiple WSNs with several gateways and cloud nodes hosting the centralized data store and application services, which also form the back end of a consumer mobile app. If the compromise is suspected to have originated at the perceptual layer (see the IoT architecture layers in Table 1) in one of the WSN segments, what sensor selection strategy should be used for analysis when dealing with hundreds of sensors? Will physical sensor location and external diagnostics capabilities, if any, allow an investigator to access the data that might be present, notwithstanding the data parsing and interpretation challenges discussed earlier? In a case where no initial pointers to the possible evidence location are available, the investigator needs to be able to correctly identify and select the elements of the IoT ecosystem from a large number of possible permutations. Incorrect selection can prevent the successful extraction of evidence or facilitate only a partial view.

IoT environments present a considerably larger number of potential individual evidence sources and introduce additional complexity due to significant technological diversity.

Aside from these unique characteristics, investigations involving the IoT will face the same fundamental jurisdictional and data ownership challenges as more traditional digital investigations involving cloud services, albeit on a much greater scale.

Table 1. The 1-2-3 Zones of Digital Forensics model and Internet of Things (IoT) architecture.[3,11,13]

Architecture layer	Layer description	Device examples	Process model zone	Applicable digital forensic areas	Evidence examples
Application	Data aggregation, storage, analytics, and dependent consumer services	Cloud services, database servers, web servers	Zone 3	Cloud forensics	Service logs, authentication data, virtual machines, and containers
Network	Communication technologies that facilitate the data transfer between layers	Gateways, firewalls, intrusion detection systems (IDS)	Zone 2	Network forensics	Packet traces, appliance logs, firewall and IDS alerts
Perceptual	A collection of heterogeneous hardware end nodes, physical sensors, and actuators	Smart appliances, mobile devices, constrained sensors, embedded readers, and tags	Zone 1	Computer, mobile, and embedded forensics	Disk images, sensor readings, routing tables, and device identifiers

IOT FORENSICS PROCESS MODELS

In response to these unique challenges, the digital forensic research community has developed several conceptual process models to guide forensic investigations involving the IoT. This effort is still in the early stages of development, with a significant focus devoted to the development of theoretical process models that are based on hypothetical case studies.[12]

The Next Best Thing (NBT) triage model was introduced in response to the challenges posed during the forensic identification phase to assist with determining the potential sources of evidence.[3] NBT recognizes the fact that devices—and any original evidence stored on them—could become unavailable or compromised due to theft, destruction, or tampering. Therefore, an investigator needs to be able to recognize other elements of the IoT ecosystem that are related to the original device in question, because these elements could contain artifacts that might have evidentiary value. The NBT principle is part of the 1-2-3 Zones of Digital Forensics process model, which can be mapped to the core three layers (perceptual, network, and application) of the IoT architecture, as shown in Table 1.

The key principle of the 1-2-3 Zones model is that zone-specific evidence extraction activities can occur in parallel as well as in isolation for cases where clear direction priorities for investigation are available. As discussed earlier and shown in Table 1, each model zone and IoT architectural layer are associated with a specific digital forensic area or set of areas. To achieve a thorough forensic investigation covering all zones, we will most likely need to apply methods

and techniques across the entire field of digital forensics. The combination of techniques from various areas of digital forensics applied at the perceptual layer has been grouped under the umbrella of device-level forensics.[13]

Similarly, the combination of techniques and resources from all digital forensic areas involved in an IoT investigation forms the conceptual construct of IoT forensics, which is used as the basis for the Forensic-Aware IoT (FAIoT) model. The key feature of FAIoT is a centralized, trusted evidence repository that incorporates a secure logging scheme, an evidence preservation module, and a provenance module, with investigator access facilitated programmatically through a read-only API. In this model, the acquisition of evidence is performed live (in real time) as part of the normal operation of a collection of IoT devices. One of the key advantages of FAIoT is the potential ability to correlate multiple types of evidence from different zones using the centralized data store. Unfortunately, the practical implications of this model and device enrollment procedures have not been tested. Sundresan Perumal and colleagues presented a more concrete top-down process model that involves significant focus on the development of specialized standard operating procedures (SOPs).[14] However, they also did not discuss the practical context of their proposed model, as it has not been practically tested.

Digital forensic readiness facilitates proactive evidence collection in anticipation of security incidents, minimizing the cost of cyber investigations.

Subsequently, Victor R. Kebande and Indrakshi Ray proposed the Digital Forensic Investigation Framework for IoT (DFIF-IoT), which focuses on establishing digital forensic readiness and increases the admissibility of evidence extracted through process concurrency.[5] Digital forensic readiness allows organizations to support digital forensic investigations by facilitating proactive evidence collection in anticipation of security incidents, thus minimizing the cost of cyber investigations. Similar to FAIoT, this model is built with IoT forensics in mind. From the readiness perspective, the model requires significant attention to proactive scenario-driven activities aimed at making sure that the environment can inherently capture the necessary evidence and implement well-defined and documented procedures as required to extract and preserve this evidence in a forensically sound manner. DFIF-IoT promotes standardization based on established practices and is modeled after the ISO/IEC 27043:2015 standard.[15]

The identification of the location of evidence can also be facilitated using the Last-on-Scene (LoS) algorithm, which states that the device that was the last node in the communication chain needs to be investigated first.[12] The LoS algorithm is applied progressively within each zone of the zone-based model, in conjunction with the NBT model, starting at the perceptual layer (Zone 1). LoS limits the scope of the investigation and decreases operational overhead by eliminating the need to examine the subsequent zones in cases where the necessary evidence has already been located in a prior zone. Nevertheless, LoS is primarily a theoretical concept and its applicability is yet to be proven in practice.

Table 2 and Figure 1 present a summary of the various process models. Arguably, the evolution of these models can be described in terms of layered growth centered around the zone-based model, which is driven by the multi-layered architecture of the IoT.

Table 2. High-level comparison of process models.

Process model	Key characteristic	Practical scope	Process coverage
1-2-3 Zones of Digital Forensics[3]	Provides a structured approach to systematically reduce complexity of investigations in IoT environments	Investigation approach mapping and assistance with identifying focus areas	Partial (artifact identification)

Next Best Thing (NBT) triage model[3]	Assists with the identification of additional potential evidence sources when primary source is unavailable	Guidance on identification of specific devices of interest within established focus areas	Partial (artifact identification)
Forensic-Aware IoT (FAIoT)[13]	Proposes to address lack of standardization in the IoT ecosystem using a centralized and secure evidence logging, preservation, and provenance service	Centralized evidence collection from heterogeneous IoT services for storage and access by investigators	Partial (artifact acquisition)
Digital Forensic Investigation Framework for IoT (DFIF-IoT) [5]	Provides a holistic approach that covers proactive (readiness) and reactive (investigation) processes in line with international standards	Improving readiness using scenario development and streamlining investigations using standardized procedures	Complete
Last-on-Scene (LoS) algorithm[5]	Assumes that the last device in the communication chain should be investigated first	Investigative guidance based on a multi-zone process flow	Complete (includes NBT triage model)

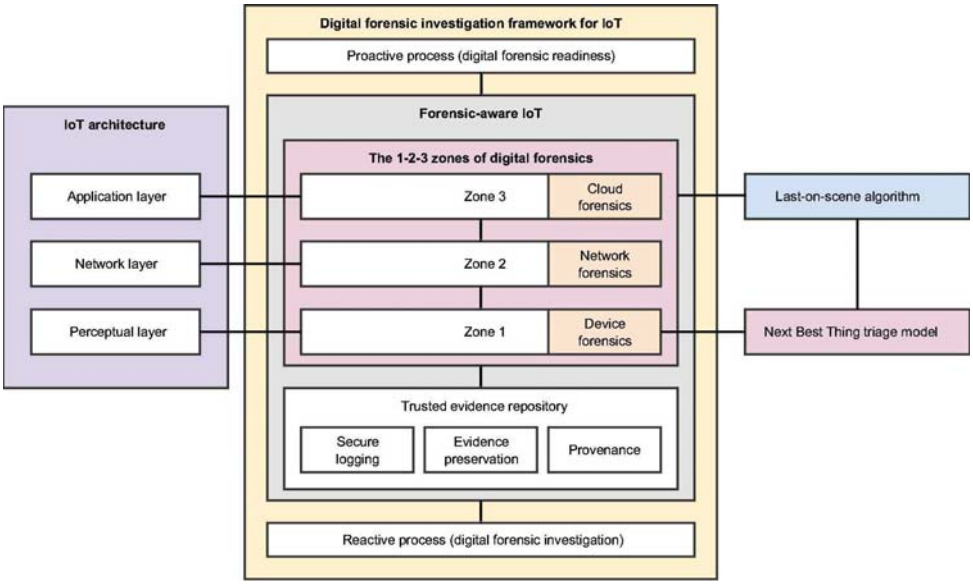


Figure 1. Conceptual digital forensic process models for the Internet of Things (IoT).[3,5,11–13]

SMART FORENSICS FOR IOT

Proactive and automated evidence collection on a large scale has the potential to address some of the challenges for the IoT paradigm.[16] In particular, automation of the forensic procedure helps reduce the operational overhead involved when dealing with a large number of potential evidence sources (such as those identified using the NBT model). Automation also increases the forensic soundness of the data-acquisition process by making it repeatable and not dependent upon manual and possibly error-prone human interactions, thus verifying the process.

To keep up with the dynamic, ubiquitous, and highly heterogeneous nature of the IoT, digital forensics needs to become "smart" without compromising adherence to the fundamental principles of acquiring admissible evidence.[16] As implied by the FAIoT model, one way to achieve a degree of smartness is to introduce real-time evidence acquisition. The acquisition does not necessarily have to be constant and ongoing, but rather could be driven by activity or anomalies. For example, IoT sensor activity detection could be based on a node's power traces using pattern recognition in power usage profiles to identify any suspicious states.[17]

The notion of real-time evidence acquisition at the perceptual layer was also conceptually explored by Nurul Huda Nik Zulkipli and colleagues.[18] In a real-time approach, the acquisition of evidence is triggered when the monitoring capability built into the processing node detects abnormal activity. However, it is unclear as to how this capability could be realized in practice, given the resource-constrained nature of typical IoT nodes and the implied need for additional computational and storage capacity. As such, the evidence acquired from IoT devices in real time would eventually be sent to the trusted repository to address the shortcomings associated with local caching on resource-constrained devices.[9] However, transfer of artifacts has the potential to increase the volume of network traffic, which could affect service availability in low-power and lossy networks such as IPv6 over Low Power Wireless Personal Area Networks (6LoWPAN).

Digital forensics needs to become "smart" without compromising adherence to the fundamental principles of acquiring admissible evidence.

In addition to real-time evidence acquisition by sensors, network traffic–based evidence can be acquired passively using specialized nodes deployed in the perceptual layer, which are commonly referred to as sniffer nodes.[19] Although passive network sniffers can capture artifacts pertaining to various network attacks such as those targeting the routing protocols, capturing all of the network traffic might be necessary and artifact acquisition from partial captures needs to be explored further.

Furthermore, artifact transfers within the same sensor network would inherit the vulnerabilities of the network. Consequently, these transfers must use strong security controls for protecting data in transit. In the case of 6LoWPAN IoT sensor networks, a common security solution is to use Datagram Transport Layer Security (DTLS) coupled with a strong encryption scheme. However, widely used SoCs such as those based on the CC2538 chip have been found to be vulnerable to DTLS implementation flaws due to inherent design shortcomings.[20] These flaws are exploitable in practice and can affect the confidentiality and integrity of evidence. As this evidence is transmitted to the trusted repository, we need additional security controls as part of the IoT system design to achieve forensic soundness.

OPEN ISSUES

To date, previously proposed process models have mostly focused on the conceptual level. Many digital forensic issues in the context of the IoT still need further investigation, some of which we address here.

Achieving Forensic Readiness

As a prerequisite for smart IoT forensics in line with the FAIoT model, real-time evidence acquisition into a trusted repository will need to be facilitated. The development of IoT products and services that can be easily integrated with remote repositories remains an open challenge. We can already see attempts to holistically address the challenge of IoT security at the state level in the US with the introduction of the proposed IoT Cybersecurity Improvement Act (2017), which targets common vulnerabilities and would make applicable mitigations more feasible in practice. Unfortunately, measures aimed at improving security do not necessarily address forensic readiness, and the integration of digital forensics readiness into IoT systems remains a challenging yet highly recommended objective.[21]

Practical Process Model, Method, and Tool Validation

Conceptual models introduced to address the unique challenges of the IoT in digital forensics are based on sound principles but still require extensive scientific validation in practice. The same applies to new methods, tools, and techniques.

Digital Warrants and Escalation

Triage models such as NBT and the LoS algorithm imply that the scope of the investigation cannot be fully determined a priori and that new potential sources of evidence will most likely be discovered during the course of the digital forensic investigation. Given the high volatility of data and the risks of evidence compromise, practical mechanisms such as digital warrants[6] are needed to enable the successful acquisition of evidence from newly identified sources (possibly remotely and on the scene).

Intelligent Evidence Analysis and Presentation

Trusted evidence repositories can aggregate a large amount of digital evidence. This evidence will likely suffer from semantic weaknesses and the lack of intelligent analysis techniques. Analyzing such data would involve correlation across heterogeneous evidence types, formats, and granularity levels to make defendable inferences based on the aggregated information.

Resolving Legal Challenges

The process models discussed in this work do not specifically address the legal challenges associated with digital forensics investigations. The legal dimension has a profound impact on successful evidence acquisition. Cross-border and multi-jurisdictional issues prevalent in the area of cloud forensics also need to be resolved in the context of the IoT, given its significant reliance on cloud-based services in the application architectural layer.

CONCLUSION

The IoT is creating new challenges for the acquisition of digital evidence, but it also has the potential to drive creation of new digital forensic techniques. As IoT-based attacks intensify and increase in frequency, successful prosecution of offenders will become ever more challenging. Current conceptual models lay the foundation for future practical work, but hands-on validation, smarter and more efficient tools, and reliable procedural guidance will be essential to conduct successful digital forensics investigations in the IoT paradigm.

REFERENCES

1. D. Evans, *The Internet of Things: How the Next Evolution of the Internet Is Changing Everything*, white paper, Cisco Internet Business Solutions Group, 2011; www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf.
2. N. Jeyanthi and R. Thandeeswaran, *Security Breaches and Threat Prevention in the Internet of Things*, IGI Global, 2017.
3. E. Oriwoh et al., "Internet of Things Forensics: Challenges and Approaches," *9th IEEE Int'l Conf. Collaborative Computing: Networking, Applications and Worksharing* (Collaboratecom), 2013, pp. 608–615.
4. C. Hauser, "In Connecticut Murder Case, a Fitbit Is a Silent Witness," *The New York Times*, 27 April 2017; www.nytimes.com/2017/04/27/nyregion/in-connecticut-murder-case-a-fitbit-is-a-silent-witness.html?mcubz=1.
5. V.R. Kebande and I. Ray, "A Generic Digital Forensic Investigation Framework for Internet of Things (IoT)," *IEEE 4th Int'l Conf. Future Internet of Things and Cloud* (FiCloud), 2016, pp. 356–362.
6. R.C. Hegarty, D.J. Lamb, and A. Attwood, "Digital Evidence Challenges in the Internet of Things," *Workshop on Digital Forensics & Incident Analysis* (WDFIA), 2014, pp. 163–172.
7. "Internet of Things Infographic," Law Enforcement Cyber Centre, 31 July 2017; www.iacpcybercenter.org/resources-2/iot/.
8. S. Watson and A. Deghantanha, "Digital Forensics: The Missing Piece of the Internet of Things Promise," *Computer Fraud & Security*, vol. 2016, no. 6, June 2016, pp. 5–8.
9. V.A. Kumar et al., "Digital Investigations for IPv6-based Wireless Sensor Networks," *Digital Investigation*, vol. 11, August 2014, pp. S66–S75.
10. J. Toldinas et al., "Suitability of the Digital Forensic Tools for Investigation of Cyber Crime in the Internet of Things and Services," *3rd Int'l Virtual Research Conf. Technical Disciplines* (RCITD), 2015; doi.org/10.18638/rcitd.2015.3.1.67.
11. A. Al-Fuqaha et al., "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, 2015, pp. 2347–2376.
12. M. Harbawi and A. Varol, "An Improved Digital Evidence Acquisition Model for the Internet of Things Forensic I: A Theoretical Framework," *5th Int'l Symp. Digital Forensics and Security* (ISDFS), 2017; doi.org/10.1109/ISDFS.2017.7916508.
13. S. Zawoad and R. Hasan, "FAIoT: Towards Building a Forensics Aware Eco System for the Internet of Things," *2015 IEEE Int'l Conf. Services Computing* (SCC), 2015, pp. 279–284.
14. S. Perumal, N.M. Norwawi, and V. Raman, "Internet of Things (IoT) Digital Forensic Investigation Model: Top-Down Forensic Approach Methodology," *5th Int'l Conf. Digital Information Processing and Communications* (ICDIPC), 2015, pp. 19–23.
15. *ISO/IEC 27043:2015 Information Technology - Security Techniques - Incident Investigation Principles and Processes*, standard ISO/IEC 27043:2015, International Organization for Standardization, 2015.
16. E. Oriwoh and G. Williams, "Internet of Things: The Argument for Smart Forensics," *The Internet of Things: Breakthroughs in Research and Practice*, IGI Global, 2017.
17. V. Looga et al., "PowerShark: IEEE 802.15.4 Mote Activity Analysis Using Power Traces and Neural Networks," *2016 IEEE Global Communications Conf.* (GLOBECOM), 2016; doi.org/10.1109/GLOCOM.2016.7842163.
18. N.H.N. Zulkipli, A. Alenezi, and G.B. Wills, "IoT Forensic: Bridging the Challenges in Digital Forensic and the Internet of Things," *2nd Int'l Conf. Internet of Things, Big Data and Security* (IoTBDS), 2017; doi.org/10.5220/0006308703150324.
19. V. Kumar, G. Oikonomou, and T. Tryfonas, "Traffic Forensics for IPv6-Based Wireless Sensor Networks and the Internet of Things," *IEEE 3rd World Forum on Internet of Things* (WF-IoT), 2016, pp. 633–638.
20. Y. Yan, E. Oswald, and T. Tryfonas, "Cryptographic Randomness on a CC2538: A Case Study," *IEEE Int'l Workshop Information Forensics and Security* (WIFS), 2016; doi.org/10.1109/WIFS.2016.7823912.
21. E. Bajramovic et al., "Forensic Readiness of Smart Buildings: Preconditions for Subsequent Cybersecurity Tests," *2016 IEEE Int'l Smart Cities Conf.* (ISC2), 2016; doi.org/10.1109/ISC2.2016.7580754.

ABOUT THE AUTHORS

Maxim Chernyshev is a researcher with Edith Cowan University. His research interests include digital forensics and wireless network security. Chernyshev received an MS in software engineering from Edith Cowan University, where he is currently pursuing his PhD. Contact him at m.chernyshev@ecu.edu.au.

Sherali Zeadally is an associate professor in the College of Communication and Information at the University of Kentucky. His research interests include cybersecurity, the Internet of Things, privacy, and networking. Zeadally received a PhD in computer science from the University of Buckingham. He is a Fellow of the British Computer Society and the Institution of Engineering Technology. Contact him at szeadally@uky.edu.

Zubair Baig is a senior research scientist in cybersecurity with CSIRO and an adjunct senior lecturer in the School of Science at Edith Cowan University. His research interests include cybersecurity, machine learning, and digital forensics. Baig received a PhD in computer science from Monash University. Contact him at z.baig@ecu.edu.au.

Andrew Woodward is a professor in the School of Science at Edith Cowan University. His research interests include digital forensics and computer network security. Woodward received a PhD in computer science from Edith Cowan University. Contact him at a.woodward@ecu.edu.au.

Experiments with Ocular Biometric Datasets

A Practitioner's Guideline

Zahid Akhtar
University of Quebec

Gautam Kumar
National Institute of Technology, Rourkela

Sambit Bakshi
National Institute of Technology, Rourkela

Hugo Proenca
University of Beira Interior

Ocular biometrics is a promising research field owing to factors such as recognition at a distance and suitability for recognition with regular RGB cameras, especially on mobile devices. The authors provide a review of ocular databases available in the literature and discuss diversities among these databases, design and parameter consideration issues during acquisition of databases, and selection of appropriate databases for experimentation.

Biometrics is a continuously evolving field that is being widely employed in applications ranging from international border crossings to unlocking smart devices. Among the various biometric characteristics (see Figure 1 and Table 1), ocular biometrics—which refers to recognizing an individual via iris, retina, sclera, periocular, or eye movements (see Figure 2)—is gaining more popularity owing to its ease of use and few user-cooperation requirements.[1]

When developing different systems based on biometric traits, experiments need to be conducted to validate the uniqueness, robustness, and feasibility of a particular trait. There are several public databases containing ocular biometric traits for researchers to experiment with—these are a vital ingredient of ongoing ocular biometrics research as they are needed in system and algorithm development, when creating a platform to be used for comparing the work of different research groups, and when introducing new challenges to the research and industry communities. Choosing the wrong dataset will produce poor results and forge the objective of the experiment, giving a false sense of progress.

To maximize the impact and usability of future ocular biometric systems, in this article we provide some guidelines for researchers and product developers to focus on choosing the proper database and evaluating ocular biometrics algorithms and systems. We also highlight open issues and challenges and discuss future research directions.

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

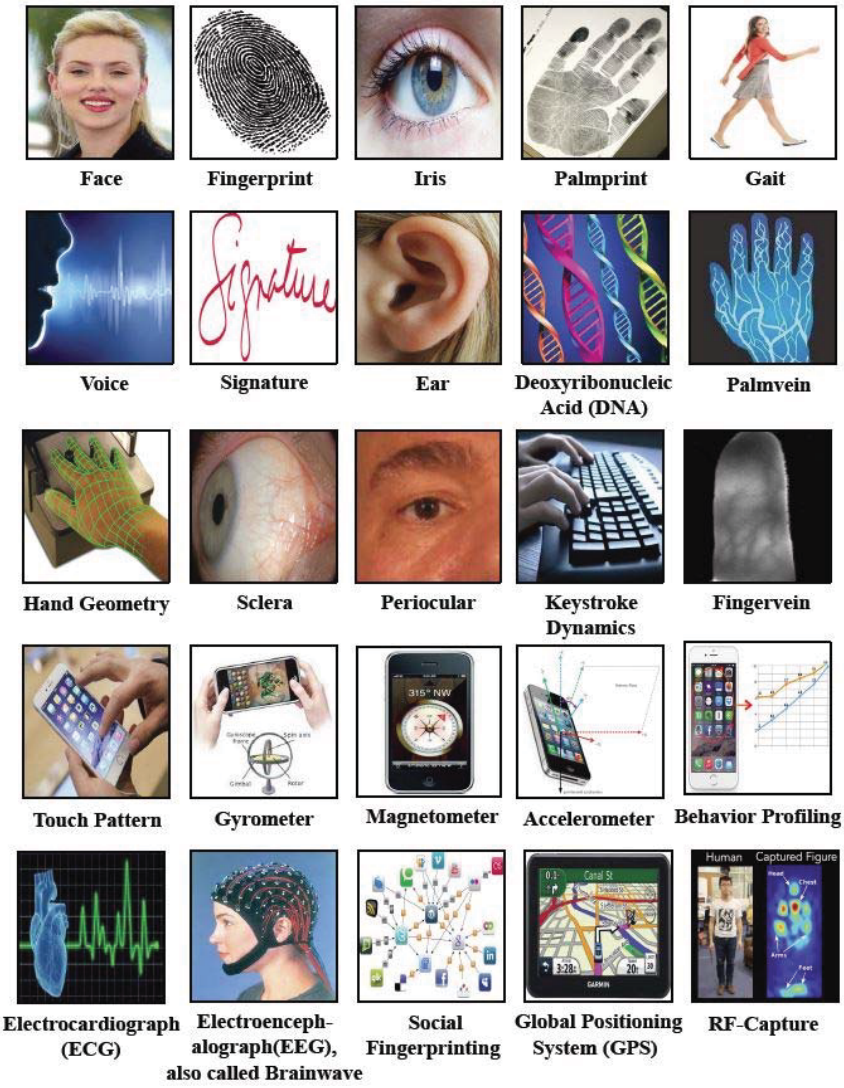


Figure 1. Examples of characteristics that have been proposed and used for person recognition.

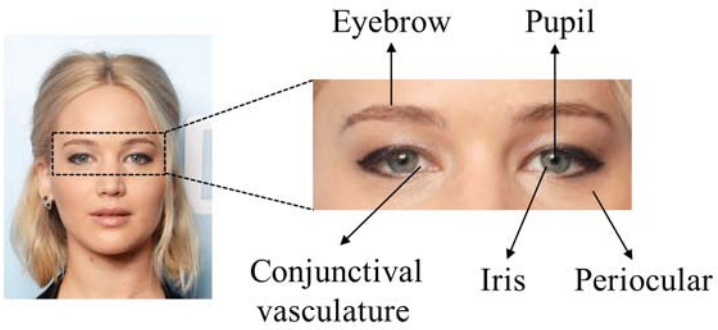


Figure 2. Ocular biometric modalities.

Table 1. Comparison of biometric traits present in the human face.

Trait	Advantages	Possible challenges
Iris	High dimensional feature can be extracted; difficult to spoof; permanence of iris; secured within eye folds; can be captured in a noninvasive way	Higher accuracy in near-infrared (NIR) images than visual spectrum (VS) images; high cost of NIR acquisition device; low recognition accuracy in unconstrained scenarios; low recognition accuracy for low-resolution images; occlusion due to use of lens; eye might close at the time of capture; does not work for keratoconus and keratitis patients
Face	Easy to acquire; yields accuracy in VS images; most available in criminal investigations	Not socially acceptable for some religions; full face template makes database large; variation with expression and age
Periocular region	Can be captured with face/iris region without extra acquisition cost	Can be occluded by spectacles; fewer features in infants
Lip	Existence of both global and local features	Difficult to acquire; less acceptable socially; shape changes with human expression

DIVERSITY IN OCULAR BIOMETRIC DATABASES

Ocular biometric databases contain different images or videos from various subjects in a maintained data structure. The data in an ocular biometric database contains the following features (usually a subset of these features).

Imaging Technique Variation

There are three types of images in an ocular biometric database:

- *Direct capture*. Samples are captured directly through sensors—usually in the visual spectrum (VS) or near infrared (NIR) spectrum—and stored in a lossless manner. Ocular recognition using different imaging modalities might result in different scores and should be reported accordingly. Tables 2, 3, and 4 represent some commonly used ocular datasets. Some sample images are shown in Figure 3.
- *Scanned capture*. Samples are scanned from printed images that have already been captured. This takes advantage of fast data processing by extracting only those parts where important information is found.[2]
- *Latent capture*. Samples are captured from some impression of the image (such as the reflection of a face in a mirror or glass).

Ocular biometric databases contain different images or videos from various subjects in a maintained data structure.

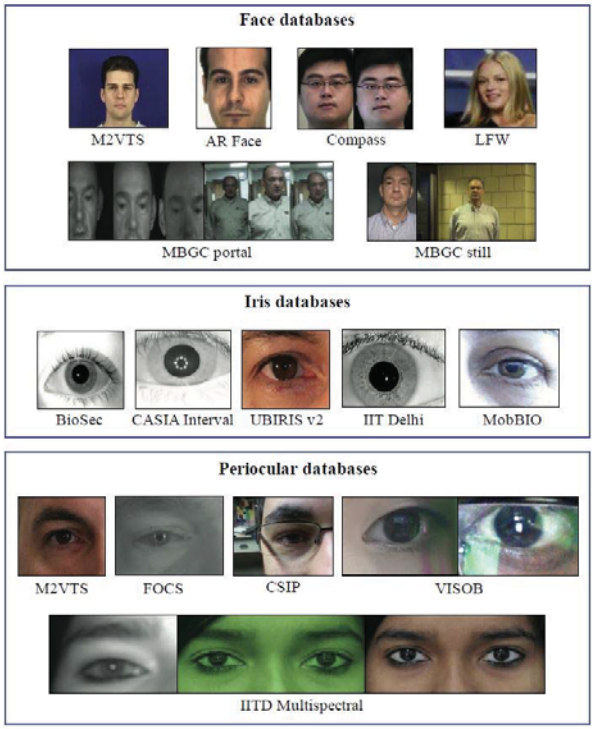


Figure 3. Samples of databases used in ocular biometric research.

Image Quality Variation

Images in the database might be of different quality, which can be obtained during data collection by changing sensor or computer-aided algorithms after data collection.

Three types of image quality variations are:

- *Spatial resolution variation*. This is the number of pixels in a unitary length—such as pixels per inch (ppi)—that mainly depends on the sensor. Higher resolution commonly leads to higher authentication accuracy.[3]
- *Bit-depth variation through bit-plane slicing*. Bit depth is color information stored in the image. Images with higher bits are expensive in terms of space, thus the bit-plane slicing method is used. Varying bit depth leads to variations in informative features of the image and accuracy.
- *Focus variation*. Change in focus produces images of varying quality such as blurred samples. Hardware and software can be used to obtain samples with varying focus properties. Techniques and standards are available for assessing the focus and quality of biometric images.[4]

Human Involvement Variation

Two types of human involvement variations are:

- *Constrained involvement*. Different impressions of the same subject can be captured by involvement of human variation in the biometric system. For example, under constrained conditions, the subject follows a mentioned expression for data collection.
- *Pseudo-unconstrained scenario*. Database images in such a scenario are acquired under uncontrolled or less-constrained environments.

Session

The time separation between two successive data-acquisition rounds is known as a session. M2VTS[5] is an example of a session-based face database. It consists of audio recordings and video sequences of 37 subjects uttering digits 0 through 9 in five sessions separated by at least one week.

Gender Specification

Gender is an important demographic attribute, which can also be used for separate recognizers to improve accuracy. Most ocular databases provide a detailed annotation of age and gender.[6]

Age Specification in Session Databases

Session databases record changes due to aging in the features of a subject over time, which can be used to improve recognition accuracy.[6]

Variation of Environment

Most databases acquired under a controlled environment facilitate the study of specific parameters on biometric recognition. However, real-time data is unconstrained in nature, where a practitioner has no control over parameters. Environmental variations largely affect the quality of acquired VS images.[7] Image acquisition location such as outdoor (cloudy/sunny day) or indoor (improper illumination) might constitute a problematic factor due to variation in illumination. BioID[8] is an example of a face database acquired in an indoor environment. It consists of 1,521 images of 23 different subjects.

Static or On-the-go Capture

Databases like UBIRIS v2[9] have distance variability, where the subject is static and standing at several distances with respect to the acquisition device/sensor. Recognition using these databases requires cooperative users, which is not often realistic. A few databases (such as MBGC)[10] consist of on-the-go acquisition images, where subjects walk through an acquisition portal.

Special Cases

Despite recent advances, there are several special challenges that still need to be solved, including identifying individuals with spectacles and identical twins. Various methods have been proposed to distinguish twins, but they require improvement for higher accuracy. Also, some diseases that affect the iris and cornea might have a negative impact on the features.[2]

CHOOSING A BIOMETRIC DATABASE FOR EXPERIMENTATION

Various ocular databases are publicly available for researchers to use for experimentation. Databases under constrained environments lack diversity, leading to low-generalization capability of systems devised using them. Databases acquired in unconstrained environments with uncooperative users (for example, operations such as recognition at a distance) contain spectacles and contact lenses, thus facilitating the capability of developing real-world robust algorithms. Databases acquired in different spectrums produce different outcomes. A researcher or practitioner should consider their research criteria and the issues mentioned here before choosing an ocular dataset. Database selection is dependent on the application—for example, for face/ocular-based uni-/multimodal recognition of moving users, one should choose a video database such as M2VTS[5] or CMU-H, whereas BioID[8] is suitable for indoor applications. For large-scale and unconstrained evaluation, Labeled Face in the Wild (LFW)[3] can be useful.

It is a very common practice by the research community to use face and iris databases for ocular recognition systems. Table 2 lists existing iris databases and Table 3 lists face databases collected in NIR and VS ranges.

Table 2. Review of existing iris databases.

Database, color model	Research lab	Version	Acquisition device	Images	Subjects	Resolution
UBIRIS (v1 RGB, v2 sRGB)	Soft Computing and Image Analysis (SOCIA) Group, Dept. of Computer Science, Univ. of Beira Interior	v1[11]	Nikon E5700	1877	241	800×600
		v2[9]	Canon EOS 5D	11102	261	400×300
CASIA (gray-scale)	Iris Recognition Research Group, Center for Biometrics and Security Research, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences	TestV1	IrisGuard AD100	10000	1000	640×480
		IRISv1	Self-developed	756	108	320×480
		IRISv2	OKI IRISPASS-h	1200	60	640×480
			CASIA-IrisCamV2	1200	60	640×480
		IRISv3-Interval	Close-up iris camera	2639	249	320x280
		IRISv3-Lamp	OKI IRISPASS-h	16212	411	640×480
		IRISv3-Twins	OKI IRISPASS-h	3183	200	640×480
		IRISv4-Interval	Close-up iris camera	2639	249	320×280
		IRISv4-Lamp	OKI IRISPASS-h	16212	411	640×480
		IRISv4-Twins	OKI IRISPASS-h	3183	200	640×480
		IRISv4-Distance	Long range iris camera	2567	142	2352×1728

		IRISv4-Thousand	Irisking IKEMB-100	20000	1000	640×480
		IRISv4-Syn	By image synthesis	10000	1000	640×480
ND-IRIS (gray-scale)	Dept. of Computer Science & Engineering, Univ. of Notre Dame	-	Iridian LG EOU2200	64980	356	640×480
MMU (gray-scale)	Multimedia Univ.	v1	LG IrisAccess2200	450	100	320×280
		v2	Panasonic, BM - ET100US, Authenticam	995	100	320×280
BATH (gray-scale)	Univ. of Bath	Iris DB 400	IrisGuard, AD-100 Dual-Eye, Autofocus Camera	8000	200	1280×960
		Iris DB 800		16000	400	1280×960
		Iris DB 1600		32000	800	1280×960
UPOL[12] (RGB)	Dept. of Computer Science, Palacky Univ. Olomouc	-	SONY DXC-950P 3CCD	384	64	576×768
BioSec (gray-scale)	Biometric Recognition Group, ATVS	-	LG IrisAccess EOU3000	3200	200	640×480
IITD[13] (Bitmap)	Biometrics Research Laboratory IIT Delhi	v1.0	JIRIS, JPC1000, digital CMOS	1120	224	320×240

MICHE (RGB)	Biometric and Image Processing Lab	v1	iPhone5, Galaxy Samsung IV, Galaxy Tablet II	1600	50	1536×2048
				1600	50	2322×4128
				1600	50	640×480
MobBIO (RGB)	Visual Computing and Machine Intelligence (VCMI), INESC TEC	-	TF300T-000128	384	105	300×200

Table 3. Review of existing face databases.

Database, color model	Research lab	Version	Images	Subjects	Resolution
FERET (RGB)	NIST	v4	14126	1191	768×512 384×256 192×128
PIE[14] (RGB)	Carnegie Mellon Univ. (CMU)	-	41368	68	3072×2048
Multi-PIE (RGB)	CMU	-	750000	337	3072×2048
SCface (grayscale and RGB)	Video Communications Laboratory, Faculty of Electrical Engineering and Computing, Univ. of Zagreb	-	4160	130	100×75 144×108 224×168 1600×1200
Yale[15] (grayscale)	Yale Univ.	-	165	15	640×480
Yale B (grayscale)	Yale Univ.	-	5850	10	640×480
ORL (grayscale)	AT&T Laboratories Cambridge	-	400	40	112×92
UMIS (grayscale)	Univ. of Manchester, Institute of Science and Technology	-	564	20	112×92
M2VTS[5] (RGB)	ACTS European Language Resource Agency	v1.0	185	37	286×350
AR[16] (RGB)	The Ohio State Univ.	-	3276	126	576×768
GTDB (JPEG)	Georgia Institute of Technology	-	750	50	640×480
Caltech (JPEG)	Computational Vision Group	-	450	27	896×592

CMU-PIE (PNG)	Vision and Autonomous Systems CMU	-	750000	337	3072×2048
FRGC (RGB, 3D channels)	Univ. of Notre Dame	-	50000	4003	1704×2272
MORPH (PGM)	Univ. of North Carolina Wilmington	-	55000	13000	400×500
PUT (JPEG)	Poznan Univ. of Technology	-	10000	100	2048×1536
Plastic Surgery (RGB)	IIIT Delhi	-	1800	900	200×200
ND-Twins (RGB)	Univ. of Notre Dame	-	24050	435	480×640
FaceExpress UBI[17] (TIFF)	Univ. of Beira Interior	-	90160	184	2056×2452
FG-NET (grayscale)	Face and Gesture Recognition Working Group	-	1002	82	400×500
CMU-H (video)	CMU	-	764	54	640×480
Compass (RGB)	CyLab Biometrics Center CMU	-	3200	40	128×128
MBGC[10] (v2 still RGB, range; v2 portal video)	NIST	v2 still	3482	437	Variable
		v2 portal	628	114	2048×2048
LFW[3] (JPEG)	Univ. of Massachusetts, Amherst	-	13233	5749	250×250

Table 4. Review of existing periocular databases.

Database, color model	Research lab	Images	Subjects	Illumination	Resolution
UBIPr[2] (RGB)	Univ. of Beira Interior	10950	261	VW	Variable
UBIPose Pr[18] (RGB)	Univ. of Beira Interior	2400	100	VW	Variable
FOCS (grayscale)	NIST Dept. of Commerce	9581	136	NIR	750×600

IMP[7] (grayscale)	Image Analysis and Biometrics Lab IIIT Delhi	620	62	NIR	640×480
		310		VW	600×300
		310		Night vision	540×260
CSIP[4] (RGB)	Soft Computing and Image Analysis Lab Univ. of Beira Interior	2004	50	VW	Variable
VISOB[19] (RGB)	Univ. of Missouri	5010381	550	VW	240×160

The number of test samples is another criterion that needs to be considered when selecting a database. For example, M2VTS[5] (which has 1,180 recordings of 295 subjects acquired over a period of four months) attracted many researchers, facilitating evaluation of many algorithms in a setup very close to real-world settings. Few databases for the periocular region such as VISOB (Visible Light Mobile Ocular Biometric)[19] are available in the public domain, as described in Table 4. As iris databases contain the eye and its immediate vicinity including eyelashes, eyelids, and nearby skin area and eyebrows, these can be used as periocular features. In turn, face databases might be cropped in a rectangular template using eye areas to be utilized as periocular datasets. Bakshi et al.[1] proposed how to optimally select a rectangular template around the periocular region.

When choosing a proper database for experimentation, a practitioner needs to know under which acquisition environment the database was captured. Next, we will discuss a typical acquisition setup and its key components. Understanding how to set up a biometric acquisition platform and what variations there are in the acquisition parameters can help a practitioner choose the right database for experimentation.

Image Acquisition Setup and Issues

Setting up an imaging environment is a critical first step to any imaging application. Figure 4 shows the image acquisition setup and parameters needed before image acquisition. Before acquiring images, the following elements and parameters need to be considered.

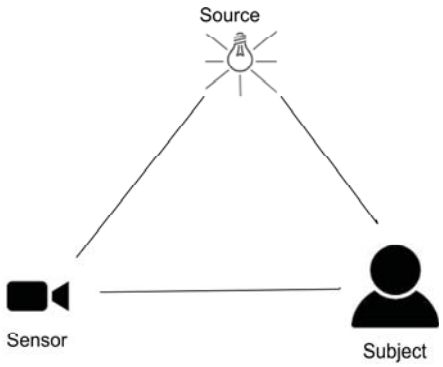


Figure 4. Image acquisition setup.

Acquisition Device Parameters

- *Imaging resolution*. The quality of an acquired image is greatly affected by resolution. High-resolution digitized images contain a wealth of features, but they require more storage space.
- *Imaging modalities*. Because VS samples suffer from illumination,[14] infrared (IR) imaging sensors are gaining much interest. The short-wave infrared (SWIR; 0.9−2.4μm) and NIR (0.7−0.9μm) spectra are reflective and eliminate indirect illumination, usually providing good image quality for recognition. SWIR and NIR spectrum databases are useful in testing cases where the application is to be done in a very controlled environment with cooperation of the subject.
- *Static or motion state*. Moving acquisition sensors usually produce blurred images and require some enhancement for feature extraction. Sometimes there is a requirement to test the performance of some method on motion-blurred images. In those cases, databases with moving cameras or objects can be considered for experimentation.
- *Focus parameter*. Setting the proper focus parameter is vital, as the wrong parameters could result in blurring of acquired image.
- *Standoff distance*. The distance between the front lens of the camera to the user under inspection is called standoff distance, which should be set according to the acquisition area of interest and the required degree of detail of the region of interest.

Lighting Setup

- *Source*. When obtaining samples with clearly visible objects, lighting conditions during image acquisition must be considered carefully. LED and lasers are good sources of light, and can reduce some illumination problems if arranged properly.
- *Characteristics of the light source*. Point light emanates concentric light and almost parallel light when placed near and far from the object, respectively. Diffuse light scatters light rays so that an object is lit from several directions. Direct light is described by rays of light following a defined direction.
- *Imaging environment*. Ambient light affects the visual appearance of objects/users, therefore the environment needs to be considered during image acquisition.

Object

- *Movement considerations*. Recognition under motion, when either the camera or the user is mobile, remains a difficult task due to blurring.
- *Constrained or unconstrained environment*. Though accuracy is higher under constrained environments, real-world applications are unconstrained, where one has no control over parameters (for example, pose).
- *Cooperative or uncooperative user*. The iris trait requires a very cooperative user and usually fails when samples are captured at a distance with low quality. Therefore, periocular recognition is gaining momentum as an alternative.

OPEN ISSUES AND FUTURE RESEARCH DIRECTIONS

Despite recent progress, several exigent problems have yet to be addressed to unleash ocular biometrics' full potential.

Heterogeneous Ocular Biometric Recognition

Cross-dataset, cross-sensor, and cross-spectral settings (in which training and testing sets are from different datasets, sensors, and spectra, respectively) are methods to assess the interoperability and generalization capability of systems. Few preliminary studies reported that ocular biometric algorithms' performance degrade remarkably under these settings. There is still room to address the

interoperability of systems under cross-settings, as this is a research direction that holds significant practical value for real-world systems.

Automatic Segmentation

Although automatic segmentation of ocular parts can help avoid those that are not beneficial (such as hair or spectacles), automatic segmentation of ocular/periocular regions is an understudied field. Reported results of automatic segmentation methods for ocular biometrics are far from the accuracy required in real-world applications, thus more attention should be placed on advanced image processing and machine learning.

Multibiometrics

It is well-documented that multimodal biometrics lead to better accuracy than the unimodal approach. However, most studies on ocular biometrics are based on a single modality. Thus, devising novel fusion schemes using ocular and other modalities needs to be explored. Further, use of image and feature quality as well as device information might be incorporated in fusion algorithms for enhanced performance. A dynamic selection-based fusion scheme might also help curb problems that arise in ocular recognition in unconstrained environments.

Webscale Ocular Biometrics

The phenomenal growth of facial and ocular videos and images on the web (in social networks and surveillance) is attracting much attention toward webscale/large-scale/open-universe biometrics. With billions of videos and images to consider, webscale ocular biometrics is a difficult task that demands speed, accuracy, and scalability. Also, there is currently no large-scale evaluation of ocular recognition schemes to establish statistical significance for published methods. Better performance might be achieved by combining meta-information associated with ocular samples. Another research track that might be pursued is formulating data-independent feature extraction and classification learning via deep neural networks.

Soft Biometrics

Soft biometrics typically refers to attributes (like gender, age, and race) that don't explicitly identify a person but complement the identity information that primary biometrics provide. Despite soft biometrics' applications in recognition, indexing, and sample retrieval, the state of the art in ocular soft biometrics is nascent, especially in unconstrained conditions. Automatic soft biometrics estimation from ocular modalities remains a challenge as demographic attributes are affected by internal and external factors, such as place of residence and worldwide cultural/racial mixing.

Ocular Biometric Spoofing and Antispoofing

Regardless of recent progress, ocular recognition systems are vulnerable to spoof attacks, which consist of submitting an artifact ocular modality, such as a replayed video of eyes, to a system. None of the existing ocular antispoofing methods exhibit low-enough error rates. One of the factors on which acceptability of ocular biometric traits depends for real-world applications is its resilience to spoofing attacks. Therefore, the biometric community should focus on devising novel measures to minimize spoofing of biometric traits. Lack of public databases containing ocular/periocular spoofing attacks has further stymied research on this topic.

Unconstrained Periocular Recognition at a Distance

Among all ocular biometric traits, periocular modality requires the least constrained acquisition process. Moreover, periocular modality can be captured at large stand-off distances (for example,

in surveillance applications) and efficiently used for personal recognition. Nonetheless, compared to other areas, periocular recognition at a distance is less analyzed.

Mobile Ocular/Periocular Recognition

The ubiquity of mobile devices with cameras has led to nearly limitless applications for ocular recognition technology. Nonetheless, mobile processing power is limited, and even commercial mobile ocular/periocular systems are either vulnerable to spoofing or produce a high level of false positives on a large dataset. Moreover, existing methods in the literature are unsuited for mobile applications because of the complex features they analyze or high computational cost. To make such applications more practical, researchers must address the issue of ocular/periocular recognition on mobile devices.

CONCLUSION

In recent years, a number of ocular biometric trait datasets have been made available to the public by different research groups. However, there is a gap between the requirements postulated by the intended biometric applications and solutions offered in many publications using these datasets. In this article, we provided guidelines for researchers and product developers to focus on choosing the right database and evaluating ocular biometrics algorithms and systems. We hope that following these guidelines will enhance the likelihood of the results obtained in a laboratory being generalized to operational scenarios.

REFERENCES

1. S. Bakshi, P.K. Sa, and B. Majhi, "Optimized Periocular Template Selection for Human Recognition," *BioMed Research Int'l*, vol. 2013, 2013; doi.org/10.1155/2013/481431.
2. C.N. Padole and H. Proenca, "Periocular Recognition: Analysis of Performance Degradation Factors," *5th IAPR Int'l Conf. Biometrics* (ICB), 2012, pp. 439–445.
3. G. Huang et al., *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*, technical report 07-49, Univ. of Massachusetts, 2007.
4. G. Santos et al., "Fusing Iris and Periocular Information for Cross-Sensor Recognition," *Pattern Recognition Letters*, vol. 57, 2015, pp. 52–59.
5. S. Pigeon and L. Vandendorpe, "The M2VTS Multimodal Face Database (Release 1.00)," *Proc. First Int'l Conf. Audio-and Video-Based Biometric Person Authentication* (AVBPA), 1997, pp. 403–409.
6. Z. Akhtar et al., "Face Recognition under Ageing Effect: A Comparative Analysis," *Int'l Conf. Image Analysis and Processing* (ICIAP), 2013, pp. 309–318.
7. A. Sharma et al., "On Cross Spectral Periocular Recognition," *IEEE Int'l Conf. Image Processing* (ICIP), 2014, pp. 5007–5011.
8. O. Jesorsky, K.J. Kirchberg, and R.W. Frischholz, "Robust Face Detection Using the Hausdorff Distance," *Int'l Conf. Audio- and Video-Based Biometric Person Authentication* (AVBPA), 2001, pp. 90–95.
9. H. Proenca et al., "The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-the-Move and At-a-Distance," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, 2010, pp. 1529–1535.
10. P.J. Phillips et al., "Overview of the Multiple Biometrics Grand Challenge," *Int'l Conf. Biometrics* (ICB), 2009, pp. 705–714.
11. H. Proenca and L.A. Alexandre, "UBIRIS: A Noisy Iris Image Database," *Int'l Conf. Image Analysis and Processing* (ICIAP), 2005, pp. 970–977.
12. M. Dobes et al., "Human Eye Localization Using the Modified Hough Transform," *Optik-Int'l J. Light and Electron Optics*, vol. 117, no. 10, 2006, pp. 468–473.
13. S. Barra et al., "Ubiquitous Iris Recognition by Means of Mobile Devices," *Pattern Recognition Letters*, vol. 57, 2015, pp. 66–73.

14. T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *Proc. Fifth IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2002; doi.org/10.1109/AFGR.2002.1004130.
15. P.N. Bellhumer, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, 1997, pp. 711–720.
16. A. Martinez and R. Benavente, *The AR Face Database*, technical report 24, CVC, 1998.
17. E. Barroso et al., "Periocular Recognition: How Much Facial Expressions Affect Performance?," *Pattern Analysis & Applications*, vol. 19, no. 2, 2016, pp. 517–530.
18. C.N. Padole and H. Proenca, "Compensating for Pose and Illumination in Unconstrained Periocular Biometrics," *Int'l J. Biometrics*, vol. 5, no. 3/4, 2013, pp. 336–359.
19. A. Rattani et al., "ICIP 2016 Competition on Mobile Ocular Biometric Recognition," *IEEE Int'l Conf. Image Processing* (ICIP), 2016, pp. 320–324.

ABOUT THE AUTHORS

Zahid Akhtar is a post-doctoral researcher with the INRS-EMT Center at the University of Quebec. His research interests include computer vision, pattern recognition, and image processing with applications in biometrics, affective computing, security systems, and multimedia quality assessment. Akhtar received a PhD in electronic and computer engineering from the University of Cagliari. He is a member of the IEEE Signal Processing Society. Contact him at zahid.eltc@gmail.com.

Gautam Kumar is pursuing a PhD in the Department of Computer Science and Engineering at the National Institute of Technology, Rourkela. His research interests include biometric security, image processing, and machine learning. Contact him at mrgautam15@gmail.com.

Sambit Bakshi is an assistant professor at the Centre for Computer Vision and Pattern Recognition in the Department of Computer Science and Engineering at the National Institute of Technology, Rourkela. His research interests include visual surveillance and biometric security. Bakshi serves as an associate editor of *Expert Systems*, *IEEE Access*, *PLOS One*, *Innovations in Systems and Software Engineering: A NASA Journal*, and *International Journal of Biometrics*. He received a PhD in computer science from the National Institute of Technology, Rourkela. Bakshi is a member of the IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence. Contact him at sambitbaksi@gmail.com.

Hugo Proenca is an associate professor in the Department of Computer Science at the University of Beira Interior. His research interests include biometrics and visual surveillance. Proenca received a PhD in informatics engineering from the University of Beira Interior. He is the coordinating editor of the *IEEE Biometrics Council Newsletter* and the area editor (ocular biometrics) of the *IEEE Biometrics Compendium*. Proenca is a member of the editorial boards of *Image and Vision Computing* and *International Journal of Biometrics* and served as guest editor of special issues of *Pattern Recognition Letters*, *Image and Vision Computing,* and *Signal, Image and Video Processing*. Contact him at hugomcp@di.ubi.pt.

The Evolving Cyberthreat to Privacy

A.J. Burns
The University of Texas at Tyler

Eric Johnson
Vanderbilt University

Cyberthreats create unique risks for organizations and individuals, especially regarding breaches of personally identifiable information (PII). However, relatively little research has examined hacking's distinct impact on privacy. The authors analyze cyber breaches of PII and found that they are significantly larger compared to other breaches, showing that past breaches are useful for predicting future breaches.

The Internet is increasingly becoming a conduit for individuals' personal and professional lives worldwide. This digital ecosystem often requires the transmission of personal information across secure and insecure networks, introducing novel information security and privacy issues and a complex chain of custody for personally identifiable information (PII). PII can be defined as "(1) any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth, mother's maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information."[1]

The Internet exposes PII to an increasing number of threats across distinct information states: data at rest, data in motion, and data in use.[2] Therefore, as organizations collect, process, and transmit more PII of customers, employees, or constituencies, they become more attractive targets for cybercriminals.[3] The massive consumer base of some online service providers presents an unprecedented opportunity for malicious parties to perpetrate large-scale security breaches. For example, Yahoo has acknowledged that a breach in 2013 might have compromised as many as one billion user accounts.[4] Even more recently, Equifax acknowledged a breach potentially affecting more than 100 million people in the US.[5] Despite the risks, consumers conduct more and more of their daily lives over the Internet, from social media to email to e-commerce. The shift to Internet-enabled transactions in the US is clear: the growth of e-commerce sales are far outpacing that of traditional sales channels.[6] This reality reflects the fact that US consumers are willing to assume privacy risks in exchange for the benefits of online engagement, based on an evaluation described as a kind of "privacy calculus."[7]

At the root of individuals' privacy-related decisions is trust.[8] However, even in the face of a data breach, many consumers' online behavior choices appear rather sticky. For example, a recent survey found that 89 percent of customers chose to maintain their relationship with an organization post-breach.[9] One explanation is that incessant reportage of so-called "mega-breaches" and annual studies warning of increased hacking activity has led consumers to disengage due to

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

"breach fatigue."[10] Whatever the cause, consumers' online choices might lead to what could be called a "non-virtuous cycle," whereby individuals continue to put more PII at risk, leading to more opportunities for breach.

In contrast to the seemingly weak market response to data breaches of individual consumers, industry reports indicate that organizations are increasingly concerned about cybersecurity.[11] In fact, cybersecurity has become a top priority for many academicians, practitioners, and lawmakers. For example, the US government announced the creation of a Cyber Threat Intelligence Integration Center in 2015.[12] Despite this increased focus on cybersecurity, the 2016 presidential election was roiled with accusations of nation-state hacking and foreign influence.[13] The implication is that cyberthreats are fundamentally changing the information security risk profile of organizations and even nations.[14] In response, the cyberthreat vector is changing the way security professionals view security risk.

In light of the complexity of cyberthreats to individuals' PII, we contend that important research remains to be done to help clarify the full extent of the problem. For example, there is a temptation to relate the prevalence of hacking with the magnitude of exposures via hacking events. That is, a natural assumption might be to expect a greater instance of hacking to result in a greater number of records breached. However, cyberthreats are asymmetric: a large number of hacking events can result in the breach of only a few records, and a single hack can expose millions upon millions of records. This results in little or no relationship between these factors. Additionally, the term "hack" has become synonymous with "breach." This is an important point because there are reasons to believe cyberthreats are unique to other breach types.[15] Despite myriad headlines and high-profile breaches, relatively little research has examined the distinct impact of the threat of hacking on privacy.

We take advantage of this opportunity and take a quantitative approach to explore cybersecurity risk. To assess the evolution and impact of cyberthreats to information security, we analyzed breaches from a well-known, trusted source of US breach data—the Privacy Rights Clearinghouse (PRC; www.privacyrights.org/data-breach)—over a five-year period. The PRC tracks losses of records containing PII in the US. As such, all breached records analyzed in this article represent a threat to privacy by including some form of PII.

Specifically, we seek to answer the following research questions.

- Do losses from cyberthreats differ from losses from other threat types?
- Are the losses from cyberthreats worsening over time?
- Are past cyberthreat breach distributions useful for predicting future distributions?
- Is a cost model derived from multi-source data useful for estimating the cost of cyberthreats?

DO LOSSES FROM CYBERTHREATS DIFFER FROM LOSSES FROM OTHER THREAT TYPES?

To help clarify the true impact of cyberthreats, we set out to establish whether hacking breaches differ from other breach types in terms of magnitude. As in the PRC dataset, hacks are breaches that result from outside party hacking or malware infection.

The first step in determining the uniqueness of the hacking threat to other breach types is to examine their distribution. The raw distribution of breaches for both hacks and other breach types is heavily skewed, with a substantial share of reported events resulting in zero and very small losses. However, the extreme values point to the unique threat of the cyber channel with a maximum loss from a single event being eight times larger than the maximum loss of any single event from all other breach types (56 million compared to 7 million).

There are a large number of breaches resulting in a loss of zero or an unknown number of records. We simplified our analyses to events that resulted in known non-zero losses. As with previous researchers analyzing the full PRC dataset (not evaluating hacks separately), we found that the breach data is approximately lognormal.[16, 17] Next, we performed an independent sample t-

test, which affirmed that losses from hacks were more severe than losses from other threat vectors [hacks: mean of ln(recs) = 8.58; other: mean of ln(recs) = 7.26; $p = 0.000$].

These results support the assertion that cyberthreats are unique to other threats. Specifically, when compared to other breach types, hacks expose significantly more records per breach. This makes sense as intelligent hackers are looking for large pools of data and scale their hacking efforts accordingly. Further, the cyber channel provides novel opportunities for thorough reconnaissance campaigns and longer-duration breach events, as malicious parties often maintain access to victim networks for extended periods. There is no comparable vector to the cyber channel in terms of potential for loss. For example, another common threat vector—lost or stolen devices—provides no such opportunity for long-term data exfiltration campaigns.

ARE THE LOSSES FROM CYBERTHREATS WORSENING OVER TIME?

Having established that cyberthreats result in larger-magnitude losses than other breach types, we next set out to determine whether the losses from the cyberthreat vector are worsening over time. To examine whether the threat has grown, we constructed a means test to examine differences in the geometric means in 2010 and 2014. We found that for the non-zero log-scaled record distributions, the geometric mean shifted from 7.63 to 10.64 between 2010 and 2014. These results provide evidence that the threat curve is shifting over time.

To help interpret this curve shift, we plotted the cumulative distribution functions (shown in Figure 1). Further illustrating the significance of the change from 2010 to 2014, we labeled the cumulative probability of a non-zero record hacking breach exposing one million records (ln(records) ≈ 13.816). In 2010, 98.52 percent of non-zero record hacks exposed one million records or fewer. In 2014, the proportion of hack-related breaches exposing one million or fewer records was down to 84.26 percent (in other words, 15.74 percent of non-zero record hacks exposed one million or more records).

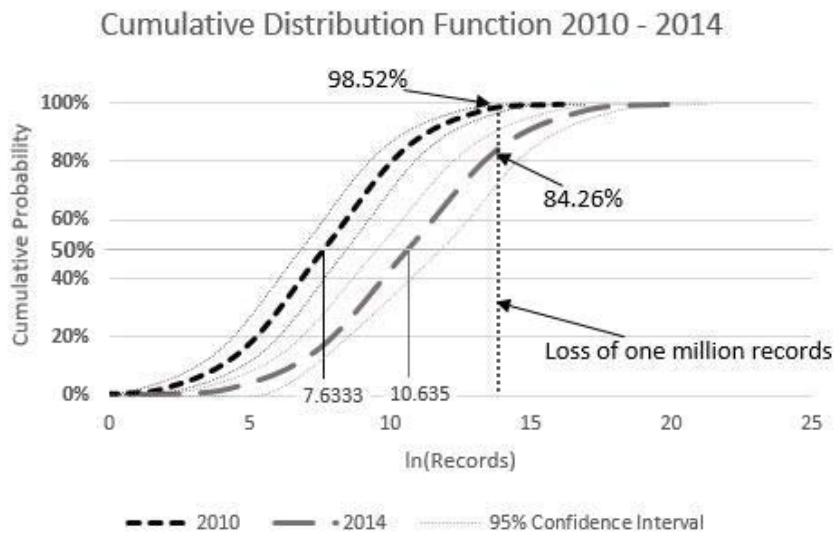


Figure 1. Hacked records per breach in 2010 and 2014. The hacking curve shifted to the right from 2010 to 2014 with means of 7.633 ln(records) and 10.635 n(records), respectively.

ARE PAST BREACH DISTRIBUTIONS USEFUL FOR PREDICTING FUTURE DISTRIBUTIONS?

Next, we set out to use our findings to predict future hacking distributions. We plotted the relationship between years and ln(records) from 2010 to 2014. To ascertain the degree to which a definable relationship explains the observed relationship between the number of records breached over time, we fitted the values to a polynomial regression equation. We found that the polynomial equation of quadratic form, $7.639 + 0.219x^2 - 0.112x$, explained 98.7 percent of the variance ($R^2 = 0.987$) in the number of records breached (log scaled), where x is the time in years since 2010 ($F = 78.34$; $p = 0.013$).

Extrapolating the linear relationship forward leads to an estimate of $e^{12.554}$ for the geometric mean of non-zero hacks for 2015. Figure 2 depicts the 2015 point estimate for the geometric mean.

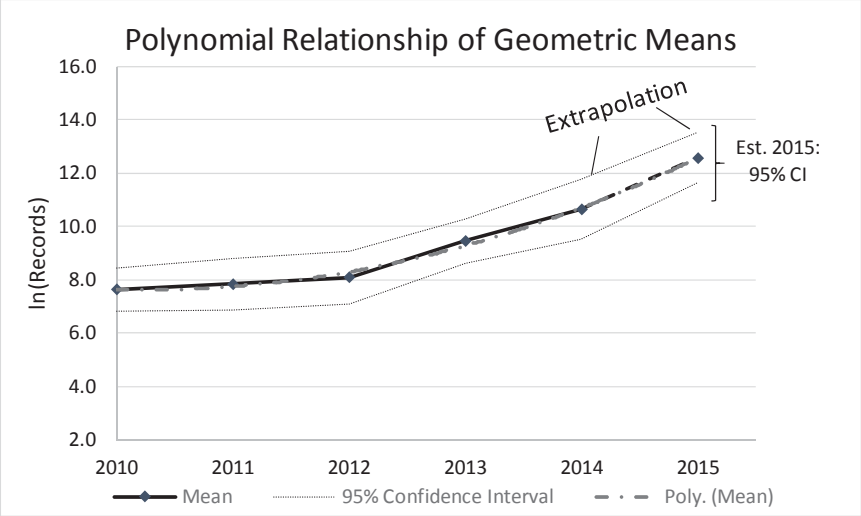


Figure 2. Polynomial relationship of breaches from 2010 to 2014 with extrapolation to 2015.

Because the variance was found to be equal across all five years from 2010 to 2014, to develop our predictive model we took the average standard deviation across these years as an estimate for the 2015 standard deviation (projected standard deviation = 3.1616) and used the average 95 percent confidence interval across these as well (projected upper limit = 13.49; projected lower limit = 11.62). With estimates for mean and standard deviation, we constructed the projected cumulative distribution function (CDF) for 2015 and compared it with the actual 2015 CDF to examine the usefulness of our approach.

Given our discovery of actual mean, standard deviation, and distributional assumption of the 2015 breach data, we constructed the predicted and actual CDFs based on our projections and the actual data. As shown in Figure 3, the projected and actual CDFs were extremely similar. In fact, the model developed using data from 2010 to 2014 provided an almost perfect prediction of the cumulative probability of losses involving greater than one million records with approximately 34 percent of non-zero record breach events containing one million or more records.

We concluded that extrapolating the identified positive relationship among years and records per breach for 2010 to 2014 produced a fairly reliable estimate of records per breach for 2015.

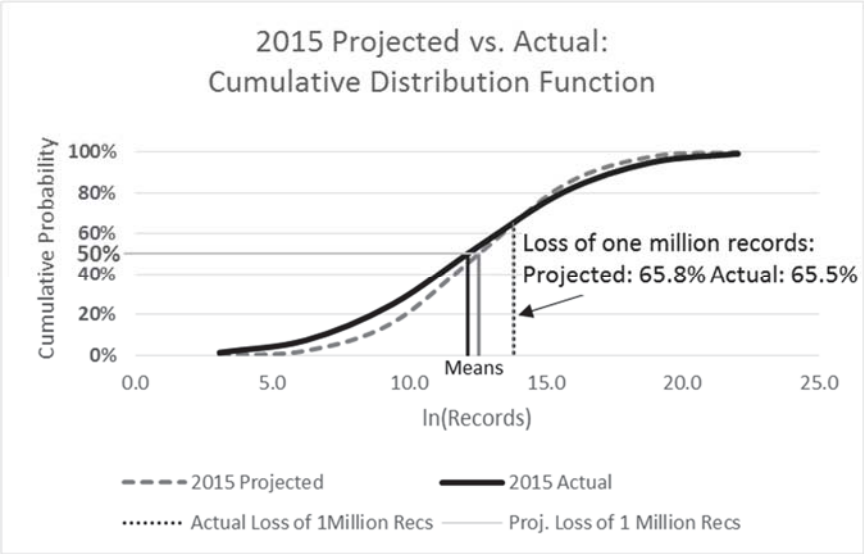


Figure 3. Projected 2015 cumulative distribution function. Projected: mean = 12.554, standard deviation = 3.162; actual: mean = 12.153, standard deviation = 4.089.

IS A COST MODEL DERIVED FROM MULTI-SOURCE DATA USEFUL FOR ESTIMATING THE COST OF CYBERTHREATS?

There have been several high-profile estimates of the cost of a data breach ranging from hundreds of dollars per record[11] to fractional dollars per record.[18] The distinction among these has largely been their generalizability across breach size. For example, Ponemon reports that their estimate of $208 per record is not applicable for data breaches over 100,000 records. However, according to our analyses, in 2015 roughly 34.5 percent of non-zero record breaches involved one million or more records.

A primary shortcoming of a dollar-per-record metric is that it ignores any economy-of-scale advantage an organization might gain from a large breach. That is, the marginal cost of additional records in response to a data breach is a decreasing function. The cost curve is particularly flat for marginal records in a very large data breach. For example, the total cost to remediate 50 million records and 50.1 million records will be more similar than the cost to remediate 100,000 records and 200,000 records despite a raw difference of 100,000 records in both instances. This is because there is a mix of fixed and variable costs associated with breach mitigation. Fixed costs are those that are relatively stable across varying numbers of records breached and include cost categories such as incident response team, public relations, and digital forensics. Variable costs are those that directly or indirectly fluctuate with the number of records lost, such as the cost to provide identity theft protection for each breach victim.

In 2015, roughly 34.5 percent of non-zero record breaches involved one million or more records.

Based on assessment of thousands of confirmed breaches and tens of thousands of security incidences, analysts at Verizon published a cost-curve estimate to help explain the true cost of breach response accounting for the breach size.[18] They found that the relationship between breach size and response cost follows a log-log linear relationship. In other words, the relationship between response cost and records breached is a power function. Table 1 and Figure 4 show the Verizon breach cost estimates.

Table 1. Verizon breach cost estimates.

Records breached		Low average	Expected value	High average
100	raw	$18,120	**$25,450**	$35,730
	ln	$9.80	**$10.14**	$10.48
1,000	raw	$52,260	**$67,480**	$87,140
	ln	$10.86	**$11.12**	$11.38
10,000	raw	$143,360	**$178,960**	$223,400
	ln	$11.87	**$12.09**	$12.32
100,000	raw	$366,500	**$474,600**	$614,600
	ln	$12.81	**$13.07**	$13.33
1 million	raw	$892,400	**$1,258,670**	$1,775,350
	ln	$13.70	**$14.05**	$14.39
10 million	raw	$2,125,900	**$3,338,020**	$5,241,300
	ln	$14.57	**$15.02**	$15.47
100 million	raw	$5,016,200	**$8,852,540**	$15,622,700
	ln	$15.43	**$16.00**	$16.56

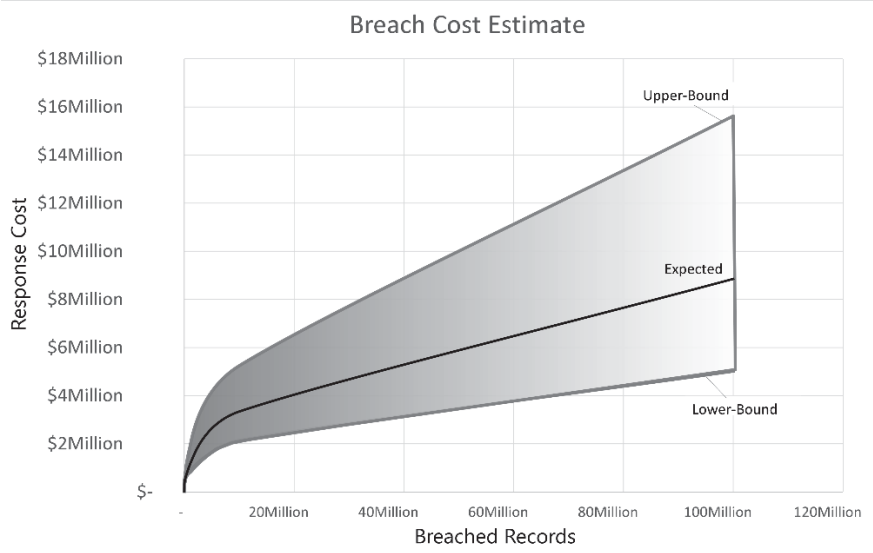


Figure 4. Verizon breach cost estimates.

Deriving the cost function shown in Figure 4 yields $y = 3618x^{0.4236}$ ($R^2 = 1.0$), where y is the cost estimate (in raw dollars) and x is the (raw) number of records breached. Therefore, the log-log relationship can also be written as the linear function $y = 0.4236x + 8.1938$, where y is ln(cost) and x is ln(records). Adding the error term to our previous cost function results in a cost model of the form $y = 0.4236x + 8.1938 + \varepsilon$, where y is ln(cost) and x is ln(records) max of $x = 18.683$; [observed overall maximum in the PRC dataset]; min of $x = 0.69$ [observed overall minimum in

the PRC dataset], and ε is the ln(cost variance) [μ = 0; σ = 0.178]. Table 2 affirms that our procedure provides cost estimates in line with the Verizon data.

Table 2. Simulated cost estimates.

Records	Size	Minimum	Mean	Maximum
100 (±0.5)	n=21	$18,564.38	$25,241.88	$32,630.41
1,000 (±1)	n=18	$53,852.28	$72,998.20	$96,374.80
10,000 (±10)	n=24	$125,847.00	$165,034.97	$252,470.13
100,000 (±100)	n=29	$286,184.29	$454,642.24	$698,069.66
1,000,000 (±1,000)	n=11	$898,846.81	$1,221,412.41	$1,584,211.05
10,000,000 (±10,0000)	n=29	$2,258,982.48	$3,313,083.69	$4,745,512.73
100,000,000 (±100,0000)	n=1	$9,241,390.96	$9,241,390.96	$9,241,390.96

Cost of Hacking

Based on our derived cost model and the CDF of non-zero record hacks, we examined the cost of hacking incidents for 2014, 2015, and our projected 2015 estimates from above. To do this, we applied a Monte Carlo approach to simulate three datasets of 100,000 observations based on our derived breach distributions and the cost model. We generated three sets of simulated breach costs in the form $y = 0.4236x_{abb'} + 8.1938 + \varepsilon$, where y is ln(cost), x is ln(records), and ε is the cost variance (μ = 0; σ = 0.178) [$x_a = 2014_{\text{actual}}$; $x_b = 2015_{\text{actual}}$; $x_{b'} = 2015_{\text{Projected}}$; max of $x_{abb'}$ = 18.683[observed overall maximum in the PRC dataset]; min of $x_{abb'}$ = 0.69[observed overall minimum in the PRC dataset].

The resulting estimate for the average cost per breached record for 2014 is $0.44, which is more in line with the Verizon average for the 2014 breach cost at $0.58 per record than other estimates that did not include breaches of magnitude greater than 100,000 records. Therefore, we believe that our computed average loss per record indicates that the distribution of records per breach drawn from the PRC dataset is roughly generalizable across the distinct Verizon breach dataset. This is a very important point because our procedure provides a way to compare the distribution of the datasets indirectly by assessing their resulting cost estimates.

In summary, by applying a cost function derived from the Verizon dataset, we produced similar cost estimates for the PRC data (for example, $0.44 per record versus $0.58 per record for 2014).

Organizations seeking to quantify risk must be able to appropriately segment cyber-risk from other information risk types.

CONCLUSION

Our investigation into the cyberthreat channel has several important implications for both researchers and practitioners. Specifically, we examined five research questions that help clarify the impact of the cyberthreat vector. First, we found that the losses from hacking are significantly higher than other breach types. Organizations seeking to quantify risk must be able to appropriately segment cyber-risk from other information risk types.[15] Second, we found that losses from hacking events appear to be worsening over time. Third, we were able to predict the higher

losses associated with cyberthreat vectors in 2015 through a relatively simple quantitative model. Finally, our cost model based on Verizon cost data produced similar cost-per-record estimates when applied to the PRC dataset.

Cyberthreats create unique risks for organizations and individuals. Attempts to quantify security risks should carefully consider the differences among risks posed by cyberthreats and other breach vectors. Our results show that although the potential and actual losses to cyberthreats seem like a study of unpredictable extremes, the fallout of cybersecurity becomes more manageable when taken in pools. It is important for researchers to continually develop and evaluate quantitative models that help explain and predict changes in the cybersecurity ecosystem.

REFERENCES

1. E. McCallister, T. Grance, and K. Scarfone, *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)*, government report SP 800-122, National Institute of Standards and Technology, 2010; https://csrc.nist.gov/publications/detail/sp/800-122/final.
2. J.B. Ard et al., "Information Behaving Badly," *Proc. 2013 New Security Paradigms Workshop* (NSPW), 2013, pp. 107–118; www.nspw.org/papers/2013/nspw2013-ard.pdf.
3. N.S. Safa, R. Von Solms, and S. Furnell, "Information Security Policy Compliance Model in Organizations," *Computers and Security*, vol. 56, no. C, 2016, pp. 70–82.
4. V. Goel and N. Perlroth, "Yahoo Says 1 Billion User Accounts Were Hacked," *The New York Times*, blog, 2016; www.nytimes.com/2016/12/14/technology/yahoo-hack.html.
5. B. Krebs, *Ayuda! (Help!) Equifax Has My Data!*, blog, Krebs on Security, September 2017; https://krebsonsecurity.com/2017/09/ayuda-help-equifax-has-my-data.
6. *Quarterly Retail E-commerce Sales: 1st Quarter 2016*, government report, US Census Bureau News, 17 May 2016; www2.census.gov/retail/releases/historical/ecomm/16q1.pdf.
7. H. Xu et al., "The Personalization Privacy Paradox: An Exploratory Study of Decision Making Process for Location-Aware Marketing," *Decision Support Systems*, vol. 51, no. 1, 2011, pp. 42–52.
8. M.E. Johnson and E. Goetz, "Embedding Information Security into the Organization," *IEEE Security & Privacy*, vol. 5, no. 3, 2007, pp. 16–24.
9. L. Ablon et al., *Consumer Attitudes Toward Data Breach Notifications and Loss of Personal Information*, RAND Corporation, 2016; doi.org/10.7249/RR1187.
10. J. Kwon and M.E. Johnson, "The Market Effect of Healthcare Security: Do Patients Care about Data Breaches?,"", *Workshop Economics of Information Security* (WEIS), 2015; www.econinfosec.org/archive/weis2015/papers/WEIS_2015_kwon.pdf.
11. *2016 Ponemon Cost of Data Breach Study*, Ponemon Institute, 2016; https://securityintelligence.com/media/2016-cost-data-breach-study.
12. *FACT SHEET: Cyber Threat Intelligence Integration Center*, government report, The White House, 25 February 2015; www.whitehouse.gov/the-press-office/2015/02/25/fact-sheet-cyber-threat-intelligence-integration-center.
13. S. Shane, "What Intelligence Agencies Concluded About the Russian Attack on the U.S. Election," *The New York Times*, blog, 6 January 2017; www.nytimes.com/2016/12/14/technology/yahoo-hack.html.
14. K.-K.R. Choo, "The Cyber Threat Landscape: Challenges and Future Research Directions," *Computers and Security*, vol. 30, no. 8, 2011, pp. 719–731.
15. R. Von Solms and J. Van Niekerk, "From Information Security to Cyber Security," *Computers and Security*, vol. 38, 2013, pp. 97–102.
16. B. Edwards, S. Hofmeyr, and S. Forrest, "Hype and Heavy Tails: A Closer Look at Data Breaches," *J. Cybersecurity*, vol. 2, no. 1, 2016, pp. 3–14.
17. C. Posey et al., "Taking Stock of Organisations' Protection of Privacy: Categorising and Assessing Threats to Personally Identifiable Information in the USA," *European J. Information Systems*, vol. 26, no. 6, 2017, pp. 585–604.

18. *2015 Data Breach Investigations Report*, report, Verizon, 2015; www.verizonenterprise.com/resources/reports/rp_data-breach-investigation-report_2015_en_xg.pdf.

ABOUT THE AUTHORS

A.J. Burns is an assistant professor at The University of Texas at Tyler. His research interests include the intersection of IT and business, with a focus on behavioral information security. Burns received a doctorate of business administration from Louisiana Tech University. Contact him at aburns@uttyler.edu.

Eric Johnson is dean of the Owen Graduate School of Management at Vanderbilt University. His research interests include the impact of IT on the extended enterprise and the security failures and economic incentives that drive identity theft. Johnson received a PhD in engineering from Stanford University. Contact him at eric.johnson@owen.vanderbilt.edu.

Understanding Privacy Violations in Big Data Systems

Jawwad A. Shamsi
National University of Computer and Emerging Sciences

Muhammad Ali Khojaye
Independent researcher

Big data systems have been instrumental in solving computational problems for business intelligence and predictive analysis. Despite this, they exhibit serious concerns for user privacy. The authors provide an overview of privacy in the context of big data, categorizing four types of existing privacy violations in big data systems and assessing the strengths and weaknesses of their protection techniques. They also provide measures that can be taken to strengthen users' privacy.

Big data applications have helped analyze and solve many data-science problems for businesses and governments alike. Governments have used big data applications to identify criminals, detect terrorist activities, and enhance citizen services. For example, in a smart city, vehicle movement can be tracked through sensors to determine volumes and patterns of traffic.[1] This information can then be linked with vehicle owner information to determine the relationships between age groups and their travel times and locations. This analysis can then be used for improved city planning.

Similarly, corporate organizations use big data to improve the customer experience, generate revenue, and provide cost-effective solutions. For example, a department store can keep track of customer spending and determine relevancy between types of products purchased at the store and their relationship to customers' age groups. The store can then focus on popular items that will increase sales.

Despite these benefits, there are increasing concerns that the information collected by government agencies and corporate organizations can lead to leakage of private and confidential information.

This article analyzes big data privacy threats and violations. Based on a review of the literature, we categorize these violations into four types. We also explain the effectiveness and limitations of existing solutions and suggest methods for improving privacy.

Published by the IEEE Computer Society
1520-9202/18/$33.00 ©2018 IEEE

BACKGROUND: UNDERSTANDING BIG DATA AND PRIVACY

From social networks to financial transactions and shopping records, a large amount of data is consistently being collected, integrated, and analyzed. Data analysis is extremely useful for forecasting and predictions, but it has also led to increased concerns about and violations of privacy.

Broadly speaking, a privacy violation is an undesired leakage, exposure, or inference of private or confidential information. Big data systems are susceptible to privacy violations primarily because of their large and continually growing datasets. As more data becomes available, a user's confidential information can be collected directly from a single source or be gathered indirectly through meticulous linking of data from multiple sources.[2]

An email provider might automatically scan emails from users to infer confidential information, and a user might agree to the provider's privacy terms without realizing that her emails are being analyzed. This is an example of a direct violation. In contrast, if data from the user's emails is linked with the user's data from another source, such as web search, then more information about the same user can be assessed. This is an example of an indirect violation.

TYPES OF PRIVACY VIOLATIONS IN BIG DATA SYSTEMS

Based on a literature survey, we categorized possible privacy violations in big data systems into four types: tracking by government, information collection by service providers, re-identification attacks, and data breaches. Table 1 summarizes these privacy violations along with examples and real-life incidents.

Table 1. Categories of privacy violations.

Type	Description	Examples/real-life incidents
Tracking by government	Governments run surveillance programs to improve security. They can collect confidential information through multiple means.	PRISM: For intelligence purposes, the US government collects data from major service providers.[3] Monitoring: City governments collect data for improved services such as traffic monitoring.[1]
Information collection by service providers	A service provider can collect and use a user's private data. Note that privacy can be violated unintentionally.	Auto scan: Email messages or posts from social network websites are scanned to display relevant advertisements. Accidental sharing of Google documents: Google accidentally shared user documents with other users.
Re-identification attacks	Individuals can be identified through correlation of big data sets.	Data correlation: Confidential information about a governor was identified by linking medical insurance records and a voter registration database.[4]

Data breaches	A data source can be hacked, leading to exposure of private data.	Ashley Madison: A dating website was hacked and confidential information was made public. Talk Talk: Personal details of almost 157,000 customers of the UK's major telecom provider were leaked. Experian: Hackers stole private data such as social security numbers and passport numbers. Target: Private credit card information was stolen through point of sales terminals at Target stores.

Tracking by Governments

Governments execute monitoring and surveillance programs for multiple reasons, including detecting traffic violations for implementing traffic laws and executing surveillance programs for enhancing national security and identifying anti-state elements. The information collected can be used to determine confidential user information.[1] Although tracking by the government is operational in many countries to offer valuable citizen services, it also carries the potential of violating citizens' privacy.

Information Collection by Service Providers

Service providers track user patterns for improved business models, leading to increased profits and enhanced user experiences. An email provider might auto-scan users' emails to display relevant advertisements. In return, users receive a free email service. For instance, Google states that it scans users' emails for "virus and spam protection, spell check, relevant search results, and features such as priority inbox and auto-detection of calendar events."[5] Similar trends also exist for other service providers such as Facebook, Twitter, and Microsoft. Service providers can utilize tracked information to generate user profiles and assess other confidential information. In agreeing to the providers' privacy terms and conditions, users often do not realize the extent of the information that can be collected and utilized to infer confidential data.[5]

Re-identification Attacks

Re-identification attacks occur in public anonymous datasets. Any information that distinguishes one person from another can be used for re-identifying anonymous data.[6] These datasets could have been published for social, personal, or research purposes.

A re-identification attack happens when the anonymity of data is compromised through the process of re-identification. An attacker could have personal or financial goals for a re-identification attack. An example is the Netflix incident in 2007, where it published anonymous data from its customers with the goal of improving its recommendation system. To protect the privacy of its customers, all the personal information was obfuscated. However, it was shown that confidential information from the Netflix dataset could partially be revealed by using the date of the rating and IMDB (the Internet Movie Database).[2] Similarly, in 2006, a re-identification attack exposed the identities of users from an anonymous dataset published by AOL.[7]

There are three different types of re-identification attacks.[8]

1. *Correlation attack*. This occurs when an adversary can correlate different datasets to obtain a more distinct and cohesive set of database records. A distinguishing feature of

a correlation attack is that information about a specific individual is not obtained; rather, it can contain sensitive information about a set of records. The attack used in the Netflix dataset is an example of a correlation attack.[2]

2. *Arbitrary identification attack*. The objective here is to relate at least one data entry in an accumulated dataset to the identity of a particular individual, with an adequate level of likelihood. This leads to learning all anonymously released information about that individual. The AOL privacy breach[7] is an example of this kind of attack, in which analyzing the anonymized dataset resulted in the identification of specific users.
3. *Target identification attack*. The objective of this type of attack is to target a specific individual. It succeeds only if it can link some dataset records to the identity of an individual, with an adequate level of likelihood. The confidential information revealed about Governor William Weld (see Table 1) is an example of such an attack.[4]

Although we have described re-identification attacks on public anonymous datasets, re-identification is also performed on information collected by governments and service providers.

Data Breaches

A data breach is defined as the compromise of security that leads to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of or access to protected data transmitted, stored, or otherwise processed.[9]

There are various causes of data breaches, ranging from insider threats to malware and misconfigured networks.[10] As mentioned in Table 1, the Ashley Madison incident is suspected to be an insider job, whereas the Target breach occurred because of malware-exploiting configuration issues in the network. Malware-based attacks and impersonation of the organization have been the biggest contributors to data breach incidents.[11] Other attacks such as SQL injection, physical theft, and privilege escalation are also possible.

The foremost requirement of enhancing user privacy is the existence of laws that can protect user privacy.

Financial systems and public datasets have been the biggest victims of data theft. This is understandable as data theft from such systems is likely to have the biggest impact.[10] Stolen data can be sold to the black market and used for fake IDs.

Similar to the context of re-identification attacks, data breach attacks can have different motives. They can be launched arbitrarily to acquire random information, or they could be initiated for a specific target to collect information on an explicit entity or organization.

Of the four types of privacy violations, the first two (tracking by government and information collection by service providers) are often protected through laws and privacy policies. However, concerns exist about the limits, restrictions, and legality of data collection and information retrieval through these means.

PRIVACY PROTECTION SOLUTIONS AND THEIR LIMITATIONS

The foremost requirement of enhancing user privacy is the existence of laws that can protect user privacy. This is followed by privacy-preserving and data-anonymity techniques. In addition, physical and security measures should be employed to prevent data theft. Note that these solutions are related. For instance, strong laws can enforce providers to implement measures to protect against data breaches.

Privacy Laws and Regulations

Laws have been instrumental in preserving privacy, as they can control government tracking and limit reading, analyzing, or publishing users' private data. Laws can also require service providers to implement proper mechanisms to ensure data secrecy and prevent data theft. Furthermore, laws can require service providers to incorporate privacy by design—building privacy mechanisms in during the process of design and development.[12]

Tangible laws to preserve and protect privacy are still evolving. National security has always been given the highest precedence and is considered the first priority. Governments collect private information with the promise of improved security[3] and enhanced services for their citizens.[1] However, this raises important privacy concerns related to the extent of private data collection and the intent of its usage.

PRISM[3] allows the US government to track and access customer data directly from service providers. Similarly, under the Phone Metadata Program, telecommunication providers deliver customer metadata to government agencies and Internet providers. In addition, governments can demand that providers bypass privacy and security measures and provide access to confidential information. In the case of *United States v. Lavabit,*[13] the US government demanded Lavabit's private keys to retrieve information about Edward Snowden. Handing the private key to the government means compromising the security of every user. In this case, the government compromised the privacy of more than 400,000 Lavabit account holders to track Snowden's account.

Laws can require service providers to implement proper mechanisms to ensure data secrecy and prevent data theft.

Although governments have been able to collect information from corporate organizations, there are a few instances where service providers have refused to share their customers' confidential information. For instance, Apple would not share confidential information with the FBI about a gunman who was involved in a shooting in California.[14]

Service providers publish their data usage and privacy policies, which grant rights to service providers for sharing user information with third parties and governments. Consequently, users have limited options when selecting service providers that safeguard their privacy. Many users remain ignorant about privacy and data-sharing policies. In response to this, the European Union (EU) has passed General Data Protection Regulation (GDPR), which regulates the control and export of data originating from the EU.[12] Violating GDPR will result in massive fines.

Cyber laws also exist to prevent data theft incidents. These laws require service providers to be more accountable for storing and collecting confidential information. In the case of data breach incidents, service providers are fined by the government and are asked to financially compensate victims.

Privacy-Preserving and Data-Anonymization Techniques

Anonymization techniques can be employed to preserve privacy. For instance, Tor provides anonymity on the Internet by routing TCP-based traffic through an overlay network consisting of volunteers.[15] Similarly, WhatsApp promises privacy protection through end-to-end encryption. Although such mechanisms are effective in improving anonymity, their usage remains limited. Furthermore, it remains unclear how much confidential information can be provided to the government upon request by law.

For public datasets, confidential information can be anonymized in several ways. A dataset has some quasi-identifiers (QIDs), which are used to identify individual data items. QID attributes are released to the public, whereas sensitive attributes are hidden and kept confidential. However, sensitive information can be identified by integrating the publicly available dataset with some external sources. Table 2 explains a few of the important techniques for providing anonymity in public datasets. These techniques are listed in the increasing order of their strength in safeguarding users' privacy. Note that increased privacy protection can reduce the utility of the

published data. Therefore, it is important to seek a good balance between privacy and data utility when assessing these data-anonymization techniques.

Table 2. Data-anonymization techniques.

Anonymity scheme	Description	Weakness/attack
K-anonymity	At least k number of redundant quasi-identifiers (QIDs) in the dataset; provides anonymity for k−1 individuals.[16]	Homogeneity attack: If sensitive information is homogenous across each record, confidentiality can be compromised. Background attack: With background knowledge about an individual, sensitive information can be identified.
L-diversity	Distribution of a sensitive attribute in each equivalence class has at least 1 "well-represented" value.[16]	Similarity attack: An adversary can determine likely possibilities of sensitive information. Skewness attack: Sensitive information can be identified in specific parts of data, as distribution of the sensitive information in the target data is significantly different than the sensitive information in the remainder of the data.
T-closeness	The frequency distribution of sensitive attributes within each equivalence class should be "close" (t-close, where *t* is a fixed threshold value) to their distribution of the sensitive attributes in the entire dataset.[17]	Lacks computational procedures to reach t-closeness with minimum data utility loss. That is, data utility loss is likely when achieving for t-closeness.
Differential privacy	Aims to limit the disclosure of sensitive data by limiting the impact of each individual in the answered query. This is achieved by adding appropriately chosen noises (for example, Laplace mechanism and geometric mechanism) to the aggregate results.[17]	Improper disclosure of the original data can cause data breaches. For instance, this technique protects individuals' privacy by adding sufficient noise to the query result; however, the original data still resides at the server, where it is vulnerable to data breaches.

Cybersecurity Measures to Prevent Data Breaches

Cyber-defense systems can employ multiple cohesive measures to prevent data breaches.[18] Physical security of servers that store and process data is of utmost importance. Espionage systems such as honeypots and prevention mechanisms such as firewalls are used to enhance data secu-

rity. Access logs and alert systems are also used to detect malicious activity. Furthermore, encryption mechanisms are employed for communication and data storage. However, despite these measures, data breach incidents continue to be frequent and impactful, and cybercriminals have been able to find new avenues for attacks.[12]

SUGGESTIONS FOR RESTRICTING PRIVACY VIOLATIONS

There are limited options in restricting privacy violations stemming from governments and service providers. These types of privacy violations appear to occur for the betterment of service, and in general, users compromise their privacy for them. However, re-identification attacks and data breach incidents can be extremely harmful.

Figure 1 illustrates a proposed model for limiting privacy violations and shows four types of privacy violations along with the role of different entities in limiting their effects. The double-sided arrows between different entities highlight greater cooperation among them, and the single-sided arrows signify the role of a specific mechanism in limiting a particular type of privacy violation.

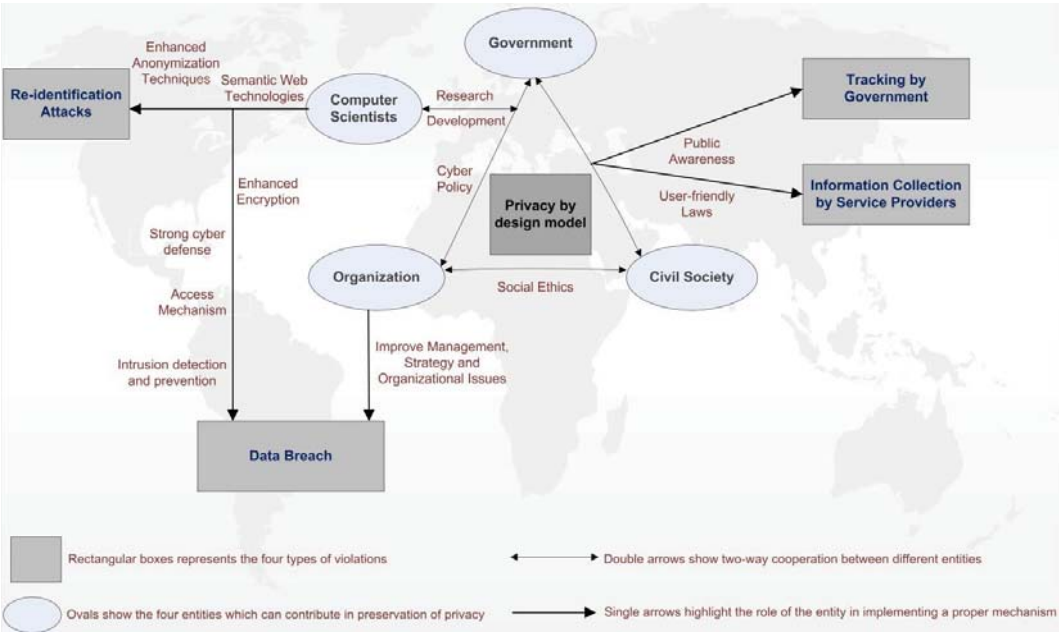


Figure 1. Model for improving privacy.

In a bigger context, public awareness about privacy is limited. The foremost requirement for avoiding privacy violations is to increase public awareness and highlight mechanisms to restrict access to confidential information. For this purpose, feedback-awareness tools can be incorporated. For instance, such tools can be useful for social networks where unwarranted photographs can be added by friends.[19] User-friendly laws can restrict collection of confidential information by service providers and governments. Laws should also be formulated to ensure privacy by design instead of privacy by choice. The roles of civil society and government are important, both in framing laws and spreading awareness. Civil society can also play a role in spreading social ethics for privacy preservation.

Strict standards are needed to ensure that confidential information is stored with enhanced security. In the case of a data breach incident, service providers should immediately inform victims. Governments should also frame laws so that organizations realize the importance of protecting private data.

The role of computer scientists is also important. They should strive to make big data storage systems more secure. This should be done by strengthening cyber-defense systems[11] and by proposing efficient methods for encryption and storage. Intrusion detection and prevention should be enhanced to limit data breach incidents. Furthermore, data access mechanisms need to be standardized to ensure authorization. This is challenging due to heterogeneity in storage mechanism for big data systems.

The research community should also focus on utilizing semantic web mechanisms such as data provenance and linkage to specify individual-specific privacy policies and to detect undesired inferences of data.[20] Anonymization techniques should be improved to reduce the effects of re-identification. Finally, governments should support research and development activities to promote innovation and standardization.

CONCLUSION

Although the importance of big data systems has been established for analytics and prediction, it is imperative that methods be adopted to preserve and protect confidential information in big data systems. Substantial and cohesive efforts are needed to achieve this important goal.

ACKNOWLEDGMENTS

This research was supported by Higher Education Commission grant HEC-NRPU 5946.

REFERENCES

1. P.A. Laplante, “Who's Afraid of Big Data?,” *IT Professional*, vol. 15, no. 5, 2013, pp. 6–7.
2. A. Narayanan and V. Shmatikov, “Robust De-Anonymization of Large Sparse Datasets,” *IEEE Symp. Security and Privacy* (SP), 2008, pp. 111–125.
3. M. De Goede, “The Politics of Privacy in the Age of Preemptive Security,” *Int'l Political Sociology*, vol. 8, no. 1, 2014, pp. 100–104.
4. D. Barth-Jones, “The ‘Re-identification’ of Governor William Weld’s Medical Information: A Critical Re-Examination of Health Data Identification Risks and Privacy Protections, Then and Now,” *SSRN*, July 2012; https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2076397.
5. M. Lindh and J. Nolin, “Information We Collect: Surveillance and Privacy in the Implementation of Google Apps for Education,” *European Educational Research J.*, vol. 15, no. 6, 2016, pp. 644–663.
6. A. Narayanan and V. Shmatikov, “Myths and Fallacies of 'Personally Identifiable Information',” *Comm. ACM*, vol. 53, no. 6, 2010, pp. 24–26.
7. M. Barbaro, T. Zeller, and S. Hansell, “A Face Is Exposed for AOL Searcher no. 4417749,” *The New York Times*, vol. 9, no. 2008, 9 August 2006; www.nytimes.com/2006/08/09/technology/09aol.html.
8. M. Jensen, “Challenges of Privacy Protection in Big Data Analytics,” *IEEE Int’l Congress on Big Data* (BigData Congress), 2013; doi.org/10.1109/BigData.Congress.2013.39.
9. *ISO/IEC 27040, Information Technology – Security Techniques – Storage*, standard ISO/IEC 27040, Int’l Organization for Standardization, 2015; www.iso.org/standard/44404.html.
10. *2016 Data Breach Investigations Report*, report, Verizon, 2016; www.verizonenterprise.com/resources/reports/rp_DBIR_2016_Report_en_xg.pdf.
11. J.A. Shamsi, S. Zeadally, and Z. Nasir, “Interventions in Cyberspace: Status and Trends,” *IT Professional*, vol. 18, no. 1, 2016, pp. 18–25.
12. M. Colesky, J.-H. Hoepman, and C. Hillen, “A Critical Analysis of Privacy Design Strategies,” *IEEE Security and Privacy Workshops* (SPW), 2016, pp. 33–40.

13. K. Zetter, "Long before the Apple FBI Battle, Lavabit Sounded a Warning," *Wired*, blog, 18 March 2016; www.wired.com/2016/03/lavabit-apple-fbi.
14. K. Zetter and B. Barrett, "Apple to FBI: You Can't Force Us to Hack the San Bernardino Phone," *Wired*, blog, 25 February 2016; Apple to FBI: You Can't Force Us to Hack the San Bernardino Phone.
15. R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second-Generation Onion Router," *Proc. 13th Conf. USENIX Security Symp.* (SSYM), 2004, p. 21.
16. A. Machanavajjhala et al., "L-diversity: Privacy Beyond k-anonymity," *ACM Trans. Knowledge Discovery from Data*, vol. 1, no. 1, 2007; doi.org/10.1145/1217299.1217302.
17. J. Soria-Comas and J. Domingo-Ferrert, "Differential Privacy via t-closeness in Data Publishing," *11th Ann. Int'l Conf. Privacy, Security and Trust* (PST), 2013; doi.org/10.1109/PST.2013.6596033.
18. S. Subashini and V. Kayitha, "A Survey on Security Issues in Service Delivery Models of Cloud Computing," *J. Network and Computer Applications*, vol. 34, no. 1, 2011; doi.org/10.1016/j.jnca.2010.07.006.
19. S. Pötzsch, "Privacy Awareness: A Means to Solve the Privacy Paradox?," *IFIP Summer School on the Future of Identity in the Information Society*, Springer, 2008.
20. C. Farkas, "Big Data Analytics: Privacy Protection using Semantic Web Technologies," *NSF Workshop on Big Data Security and Privacy*, 2014; http://csi.utdallas.edu/events/NSF/papers/paper06.pdf.

ABOUT THE AUTHORS

Jawwad A. Shamsi is a professor and head of the Department of Computer Science at National University of Computer and Emerging Sciences. His research interests include big data systems, privacy, security, cloud computing, and high-performance computing. Shamsi received a PhD in computer science from Wayne State University. His research has been funded by the Higher Education Commission of Pakistan and NVIDIA Corporation. Contact him at jawwad.shamsi@nu.edu.pk.

Muhammad Ali Khojaye is an independent author, researcher, and consultant, specializing in architecture and designing large-scale distributed applications. He is an enthusiastic and passionate technical leader with more than 10 years of commercial software development experience. Khojaye's research interests include distributed systems, particularly big data challenges and high-performance computing. He received a Master's degree in computer science from the University of Leicester. Contact him at muhammadkhojaye@gmail.com.

Senior Database Analyst

hibu Inc. is seeking a Senior Database Analyst position for its King of Prussia, PA office. The Senior Database Analyst is responsible for performing database systems development, contributes to the data management efforts in areas including data analysis, data modeling, data flows and support. This includes hands-on role in delivering logical/physical data modeling, detailed data analysis, defining data dictionary and metadata management, database design and implementation. Senior Database Analyst participates or helps support end-to-end business processes and workflows included, but not limited to, understanding and gathering source system data, processing logic, content and operational system usage, and data management analysis & development. Bachelor's Degree in Computer Science, Software Engineering, Computer Engineering, Computing Technology, MIS, Information Technology, Business Administration required. 5 years of experience in the following required: Design and Development of ETL components, Data Warehouse concepts, Data Modeling and Data Integration tools, Oracle11g RAC, Relational DB Concepts and use of complex SQL DML/DDL, Stored Procedures, Triggers. 2 yrs of experience in Salesforce.com data integration required. Send application materials to Steve Clifford at: stephen.clifford@hibu.com