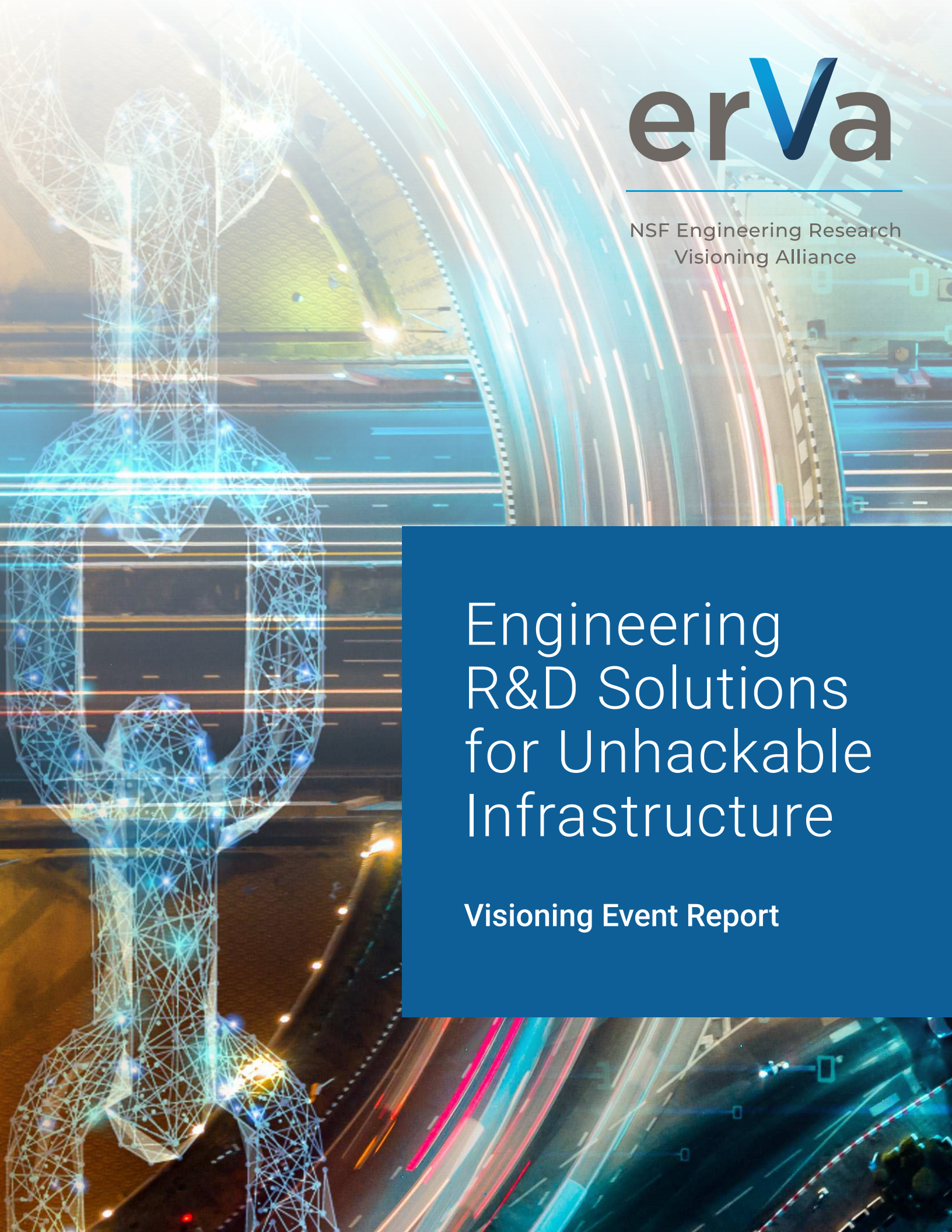


erVa

NSF Engineering Research Visioning Alliance

Engineering R&D Solutions for Unhackable Infrastructure

Visioning Event Report

Engineering R&D Solutions for Unhackable Infrastructure

A Visioning Report of the Engineering Research Visioning Alliance

Report Finalized December 1, 2022

Based on proceedings from an event hosted by:

This material is based upon work supported by the National Science Foundation (NSF) under award # 2048419. Any opinions, findings, interpretations, conclusions, or recommendations expressed in these materials are those of its authors and do not represent the views of the UIDP Board of Directors, UIDP's membership or the National Science Foundation.

The Engineering Research Visioning Alliance (ERVA) is a neutral convener that helps identify and develop bold and transformative new engineering research directions, directly supporting the nation's ability to compete in a rapidly changing global economy. Funded by the National Science Foundation (NSF) Directorate for Engineering, ERVA is a diverse, inclusive and engaged partnership that enables an array of voices to impact national engineering research priorities. The five-year initiative convenes, catalyzes and empowers the engineering community to identify nascent opportunities and priorities for engineering-led innovative, high-impact, cross-domain research that addresses national, global, and societal needs. Learn more at ervacommunity.org.

Suggested citation: Engineering Research Visioning Alliance. 2022. Engineering R&D Solutions for Unhackable Infrastructure: A Visioning Report. Columbia, SC: SSRN. ssrn.com/abstract=4294220.

Acknowledgments

ERVA is grateful to our partners at the National Science Foundation (NSF) Engineering Directorate for their ongoing engagement and support for this work, particularly Program Director Louise R. Howe and NSF Engineering Directorate leadership Susan Margulies, Don Millard, and Sohi Rastegar.

ERVA visioning events enable the engineering research community to identify nascent opportunities and priorities for engineering-led, innovative, high-impact research that addresses global and societal needs. Each event relies on the efforts of organizations and individuals who volunteer to lead, guide, and participate in its activities.

Gratitude is extended to the event co-host, the Massachusetts Institute of Technology's Industrial Liaison Program, for providing an inspiring location for ERVA's first in-person visioning event, as well as leadership and recommendations for participants. We are especially grateful to Saurabh Amin, MIT associate professor and director of the Pierce Laboratory for Infrastructure Science & Engineering, who served both as the event's Thematic Task Force co-chair and our host lead. Co-leading the Thematic Task Force was David Ott, senior researcher at VMware, who worked closely with the ERVA team for report writing and development. These visionary subject matter experts developed the creative framework for this visioning event and contributed to the report. Special thanks to report writer Susan Lang, director of the Center for the Study and Teaching of Writing at The Ohio State University, for attending the event and writing the report.

Thematic Task Force members contributed their time and expertise and were instrumental in framing and shepherding the workshop conversations:

- Kevin Fu, University of Michigan
- Sean Guillory, Booz Allen Hamilton
- Jennie S. Hwang, H-Technologies Group
- Todd Jones, Sandia National Laboratories
- Ashley Podhradsky, Dakota State University
- Zhihua Qu, University of Central Florida
- Vipin Swarup, MITRE

This visioning effort also benefitted from the expertise of facilitators in each of the breakout sessions:

1. **Human-Technology Interface Considerations:** Sean Guillory, Booz Allen Hamilton, and Ashley Podhradsky, Dakota State University
2. **Measuring and Verifying Security (Metrics):** Todd Jones, Sandia National Laboratory
3. **Future Approaches to Autonomous Security:** David Ott, VMware, and Saurabh Amin, Massachusetts Institute of Technology

4. **New Approaches to Resilience in Interdependent Infrastructures:** Zhihua Qu, University of Central Florida, and Vipin Swarup, MITRE Corporation
5. **Architecting Trustworthy Systems:** Kevin Fu, University of Michigan

In addition to a presentation from the Thematic Task Force co-chairs, the event was informed by a panel discussion moderated by Saurabh Amin, MIT, and featuring expertise from Vipin Swarup, MITRE; David Clark, MIT; and David Ott, VMware.

ERVA is also grateful to all the workshop participants who contributed their time to create a valuable discussion and visioning report and to their organizations for the liberty to share their expertise for this effort.

Kristina Thorsell, ERVA projects advisor, facilitated the event and the visioning exercise. The event and report development were executed under the guidance of ERVA principal investigator Dorota Grejner-Brzezinska, The Ohio State University, and co-Principal Investigators Anthony Boccanfuso, UIDP; Charles Johnson-Bey, Booz Allen Hamilton; and Edl Schamiloglu, University of New Mexico. Development of the visioning session theme was informed by input from the ERVA Standing Council, Advisory Board, and NSF Engineering collaborators.

Finally, we are grateful for the team that provided operational support for the event and report, representing both ERVA and UIDP (ERVA's administrative core partner): Josh Aebischer, ERVA program specialist; Michael Brizek, UIDP program director; Natoshia Goines, UIDP events manager; Jessica Hawke, UIDP events coordinator; Sandy Mau, ERVA communications director; Mark McGill, ERVA program coordinator; and Annie Shealy, ERVA events and engagement manager.

Credit: Texas Advanced Computing Center

Executive Summary

Infrastructures include at least four key components that converge to deliver useful functionality: physical infrastructure, software and algorithms, data communications, and humans. The latter design, deploy, and operate infrastructure. They can also become attack agents and threats.

The goal of "unhackable infrastructure" is for the engineering research and development community to dramatically raise the bar on robust protections against an enormous array of adversarial threats across an enormous array of infrastructure domains, many of which will be new in the days ahead.

As technological advances provide more sophisticated tools and systems (e.g., artificial intelligence-driven autonomous agents, real-time automation), adversaries will be able to leverage these new technologies to expand their capabilities. This should be anticipated by the research community.

Creating a future with unhackable infrastructure was the challenge presented to a diverse array of experts and practitioners convened by the Engineering Research Visioning Alliance (ERVA). A Thematic Task Force with members from academia, industry, and the nonprofit sectors identified five key areas for research investment to advance progress toward this aim. These are listed below, along with select key ideas emerging from the event:

01

Human-Technology Interface Considerations

- Humans are both the weakest link in security and the greatest opportunity to advance protections in cyber-physical infrastructure.
- Research should greatly increase focus on the human element within tomorrow's infrastructure; for example, including humans in system security modeling and addressing the incentives and the economics surrounding adversarial behavior and security design.

02

Measuring and Verifying Security (Metrics)

- Advances in measurement tools and metrics for tomorrow's infrastructures are sorely needed to support security risk evaluation, design tradeoffs, verification, automation, and more.
- Additional tools are needed to support observability, autonomous attack recovery, and designed-in verification techniques.

03

Future Approaches to Autonomous Security

- Autonomous security (self-configuring, self-guiding, self-managing, self-learning, self-tuning, etc.) is needed to address both complexity and scale in tomorrow's infrastructures.
- Autonomous agents addressing security functionality could be used in service of, or in synergy with, human decision-makers.
- Autonomous systems need better contextual understanding and response capabilities.

04

New Approaches to Resilience in Interdependent Infrastructures

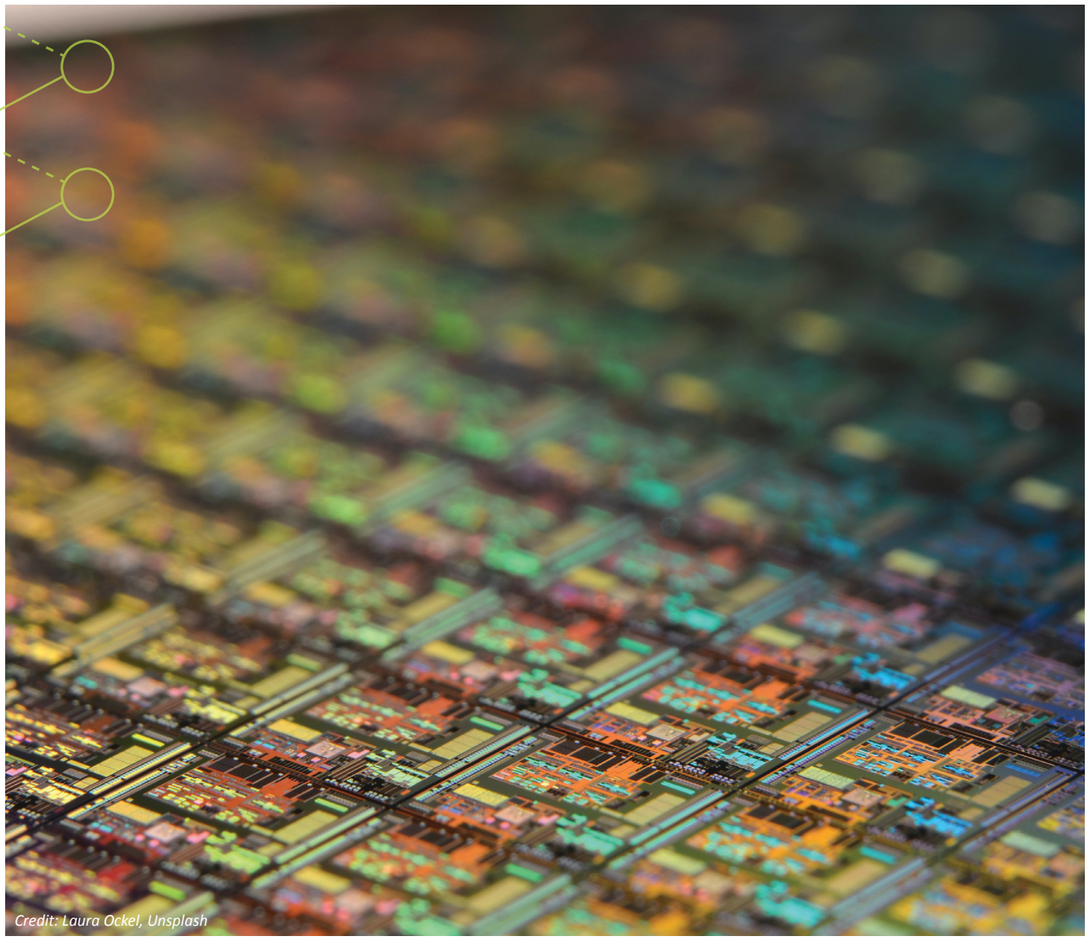
- Attackers are becoming more adept at exploiting the vulnerabilities of hierarchically managed infrastructures with tightly coupled processes and complex dependencies.
- An important opportunity for future research is to develop design approaches that maintain system-level properties of safety and security after the integration of modular components.

05

Architecting Trustworthy Systems

- Research in trustworthy systems architectures looks at how correctness of operation can be verified more robustly and built into system architectures and infrastructure.
- Some key issues include design specification, decentralized control, confidential computing, and new infrastructure domains.

The research priorities detailed in this report lay the groundwork for a future with unhackable infrastructure.



Credit: Laura Ockel, Unsplash

Taking Action

Anticipating and preparing for security threats, whether short-, medium-, or long-term, has escalated as a critical national and international priority. Engineering-led research and innovation in these priority areas is essential to lay the foundation for the high-impact, multidisciplinary research necessary to mitigate the economic and security threats inherent in cybercrime and secure both physical and virtual spaces.

Every focus area selected by the Thematic Task Force provided opportunities to progress toward unhackable infrastructure through proactive and comprehensive research approaches. The aim of this report is to inspire researchers and sponsors (public, private, and nonprofit) to pursue these priorities. ERVA challenges readers to disseminate this report and prioritize areas with potential for the greatest return on investment.

Engineering R&D Solutions for Unhackable Infrastructure

Launched with funding from the National Science Foundation (NSF) in April 2021, the Engineering Research Visioning Alliance (ERVA) provides a neutral platform to identify bold and societally impactful engineering research directions that will place the United States in a leading position to realize a better future for all. It is an engaged, inclusive, multilayered partnership, providing a truly diverse array of voices with the opportunity to impact national research priorities.

Hosted by the Massachusetts Institute of Technology on August 10-11, 2022, ERVA's visioning event on "Engineering R&D Solutions for Unhackable Infrastructure" convened 35 researchers and technical experts from academia, government, and industry to identify engineering research priorities leading to significant advancement in the security and resilience of tomorrow's cyber-physical infrastructures. The visioning exercise addresses longstanding trends in national and global scale cyberattacks on our increasingly complex and cyber-enabled infrastructures.

Cybercrime has been identified as a growing problem for decades, and the global COVID pandemic only intensified the situation. As many Americans moved to remote work, cybercriminals seized opportunities to attack. The Center for Strategic and International Studies (CSIS) list of significant cyberattacks (those on government agencies, defense, and high-tech companies, or economic crimes with losses of more than a million dollars) notes 81 such attacks from January to July 2022 and over 350 since January 2020.[1] The cost of cybercrime is predicted to top $10.5 trillion by 2025, and the destructive impacts of cybercrime were partially responsible for the Doomsday Clock moving 20 seconds closer to midnight in 2020, where it remains in 2022.[2]

Against this backdrop, visioning event participants considered the question of how to move research in infrastructure security and resilience from being merely *reactive* to being *anticipatory*—studying the challenges for future infrastructure and not just solutions for the infrastructure problems of today. Critical infrastructure systems must protect human safety and well-being and adapt to unexpected changes. These systems must actively resist adversaries (both known and unknown) and demonstrate resilience in the face of systemic risks. Developing context-driven criteria and pathways to advance these aspirations can have far-reaching impact on how future infrastructures are built, operated, and maintained, and in particular support the creation of a robust security ecosystem.

Visioning event participants considered the question of how to move research in infrastructure security and resilience from being merely reactive to being anticipatory.

To prevent potentially catastrophic cyberattacks on the nation's infrastructure systems, participants considered gaps in today's security technologies and formulated bold, new ideas and visions that could steer future research toward areas of much-needed innovation in computer security as we know it with significant return on investment. Visioning event participants asked what tomorrow's "unhackable infrastructure" could look like with non-incremental advances in security R&D and engineering.

While developing a more complete understanding of "unhackable infrastructure" was one aspect of the visioning challenge, a working view of "infrastructure" was defined as including at least four key components that converge to deliver modern cyber-physical services:

- physical infrastructure (assets, hardware),
- software and algorithms,
- data and communication networks, and
- humans who are involved at multiple levels (e.g., users, operators, security administrators, adversarial elements).

This combination makes modern and future infrastructures highly complex sociotechnical systems.

To truly be unhackable (or as close to it as possible), infrastructure must be robust against an enormous array of attacks/problems that could occur and must have the accompanying solutions to deploy countermeasures, including immediately identifying the source of a threat. System interactions with more "hackable" humans and legacy systems are another consideration. Additionally, an unhackable infrastructure would need to understand human behavior intimately to anticipate interactions. Unhackable infrastructures will enable responsible and safe operation, system-wide situational awareness, and quick recovery from disruptions. They will also promote new business models to advance the security ecosystem and shape market competition in a direction where service providers must offer security and resiliency guarantees to remain competitive.

The systems and infrastructures of tomorrow will be under continuous attack by increasingly sophisticated adversaries who will leverage, among other things, artificial intelligence (AI) techniques in the new generations of malware they create. Finding more effective and advanced engineering solutions to prepare for and respond to attacks is critical to both current and future societal-scale infrastructure systems and services.

Persistent security threats and a lack of comprehensive and visionary approaches for protecting infrastructure represent two of the main engineering challenges of our times.

Credit: Evan Dougherty/Assistant Multimedia Editor, University of Michigan, College of Engineering

Visioning Event Purpose and Structure

The overarching theme of unhackable infrastructure was selected with broad community input and due diligence performed by the ERVA Standing Council. Participants were identified and invited based on their research and expertise and included engineers and scientists across academic disciplines, geographic location, organization sector and type, gender, race/ethnicity, and career stage. The task at hand: identify unexplored or underexplored areas where additional research can make bold, positive impacts. The ERVA visioning event was structured to spark new research directions and catalyze engineering research for a more secure and resilient world.

To support the technical aspects of this effort, a strategic group of experts (Thematic Task Force) was assembled to identify key areas of research needs and to frame the event discussion. The group identified five distinct areas which became discussion themes — areas for participants to envision bold ideas, probe unexamined questions, and focus their combined expertise. Participants were placed into breakout groups based on these themes and on their scientific domain to facilitate interdisciplinary discussions around needed fundamental research in the identified areas.

Breakout Theme Topics

1. Human-Technology Interface Considerations
2. Measuring and Verifying Security (Metrics)
3. Future Approaches to Autonomous Security
4. New Approaches to Resilience in Interdependent Infrastructures
5. Architecting Trustworthy Systems

The groups also debated which of these ideas are engineering solutions that could apply to a variety of challenges versus which of these could be combined to comprise one specific challenge for which diverse engineering solutions could be developed. Following workshop breakout group reporting, a large number of issues and ideas were distilled into this report, which provides a catalog of group recommendations on research and innovation directions for the engineering R&D community.

Engineering Research Opportunities to Enable Unhackable Infrastructure

#1

Human-Technology Interface Considerations

SUMMARY OF RESEARCH PRIORITIES

Extensive work on human incentives and the economics of security and resilience is needed, including organizational and adversarial perspectives.

Usability research to address unwanted tradeoffs with functionality, convenience, cost, and more is needed.

Work to incorporate frontier user interface technologies like augmented and virtual reality (AR/VR), continuous biometric monitoring, natural language processing (NLP), and even brain-computer interfaces (BCI) into security and resilience interface design is required.

Immersive human-computer interaction environments of the future will need extensive research on threat modeling, vulnerability mitigation, and more.

A key observation in engineering research and development for unhackable infrastructures is that future research must include more attention to the role of humans in security and resilience solutions. Humans will design, deploy, and manage the cyber-physical systems of tomorrow. They will also be the users, operators, maintainers, and evaluators. Conversely, they will be the adversaries who engineer hacks, intrusions, abuses, and attacks against tomorrow's infrastructures. Thus, humans are simultaneously the weakest links in security yet the greatest opportunity to protect cyber-physical infrastructure.

Indeed, the Thematic Task Force recommends that the entire notion of Cyber-Physical Systems (CPS) (infrastructure) be recast as the challenge of building secure and resilient Cyber-Physical-Human Systems (CPHS) to more fully comprehend the role of humans in the picture. For example, research on CPHS and architecture should more explicitly model the human element within a security design, including minimally the user, the system administrator, and the adversary. Well-designed systems must be *adaptive* to the range of human behavior and unexpected inputs, *sufficiently intelligent* to recognize the difference between abnormal normality and threatening human behavior, and *sufficiently resilient* to withstand the inevitable attacks engineered by humans.

Economic Context and Incentives

MIT's David Clark framed this discussion by asking whether research on infrastructure security is overinvested in technology and underinvested in the science of human incentives. Many adversarial actions represent unanticipated uses of a technology that are still entirely *within* the system's functional specifications. **Research is needed to better understand the dynamics of how humans use and abuse technology, and how human incentives can more systematically be anticipated and guided within a secure and resilient design.**

The economics surrounding system security engineering must also be considered. Infrastructure owners and developers often lack the incentives needed to invest in robust security technologies within their infrastructures. Unlike additional features or functionality, customers simply expect security and are not accustomed to paying more for it. Research is needed to create economic incentives for technology development and investment. To illustrate the relationship, consider how system modularity and open architecture frameworks can encourage an ecosystem of innovative solution providers and fuel economic incentives. Future research may explore how engineered solutions can likewise comprehend and build in economic incentives for security and resilience.

Usability

Another human-related issue in security engineering for unhackable infrastructures is usability. Security features are often introduced at the expense of usability, adding inconvenience, limiting functionality, and/or adding cost to a system or infrastructure solution. Research is needed on innovative approaches to security and resilience that avoid or minimize such tradeoffs and to better comprehend human usability issues for various user roles (e.g., operator, administrator, auditor) within a CPHS infrastructure.

One challenge is advancing the state-of-the-art in security interface design for infrastructure operators. Research is needed to explore and enable the use of augmented and virtual reality (AR/VR) specifically for security systems, allowing operators to better visualize adversarial activity, identify attack surfaces, configure and verify defenses, build robust configurations, and more. AR/VR could be used in risk mitigation, threat modeling, and design specification. Future AR/VR could feature much-improved headsets and incorporate tactile elements to enrich operator experience and enable a more nuanced repertoire of user actions.

While authentication and authorization schemes have received considerable attention from the research community over the years (e.g., biometric forms), more work is needed to enable usability and truly revolutionary change in practice. Tomorrow's authentication should avoid widely discussed shortcomings of today's password schemes and even multi-factor frameworks by being simultaneously less intrusive and more robust. Such schemes should automatically and adaptively leverage user contextual clues and prior history to understand what criteria to apply in user authentication and authorization. Tomorrow's frameworks must perform *continuous* biometric monitoring and use a combination of technologies.

Finally, advances in natural language processing and smart personal assistants (e.g., Siri, Alexa) create new opportunities to develop the notion of a *personal security assistant*. Research may explore how to develop, support, secure, and integrate a personal security assistant or agent that is both trustworthy and likable to human users and make it platform- and application-agnostic. The range of functions that such an agent could perform for human users who need to navigate security decisions at various levels also should be explored; for example, such assistants might verify application interfaces and user inputs.

New Technical Domains

Human users of the future will leverage new modes of computer interaction, and research is needed to understand how these technologies will work, how they can be secured, and how they can be used by unhackable infrastructure designers and operators.

Brain-computer interfaces (BCI) would seem a futuristic technology for tomorrow's infrastructures but could well become a powerful mode of user input or communication. Tomorrow's work uniforms and clothing may be instrumented with technology elements that are integrated into infrastructure architectures, automatically triggering user interface elements, managing a user's identity, enabling infrastructure actions, and more. A version of this would embed additional capabilities into clothing—controllers, visualizers, verifiers, configuration devices, and more.

Human immersion in VR-enabled virtual worlds (e.g., Meta's notion of a collective "metaverse") is another direction worthy of study. Research questions may include how humans collaborate and interact within the context of tomorrow's CPHS infrastructures; what tomorrow's threats look like in this context; identification of potential mitigations; and how such environments could enable human decision-making.

#2

Measuring and Verifying Security (Metrics)

SUMMARY OF RESEARCH PRIORITIES

Challenges in measuring, evaluating, and verifying security in CPHS are considerable because of their inherent complexity, dynamic and uncertain environments of use, constituent human elements, and more.

Continuous monitoring and automated response research at CPHS interfaces in the context of changing threat landscapes and unpredictability of overall system behavior is needed.

Because observability is a key design issue in future systems, **foundational research and practical tools are needed to observe, estimate, and update the dynamic security state of a CPHS**, and to use it for improved situational awareness, cyber deception, and adversarial response.

Fully automated mechanisms are needed to maintain key functionality (resilience) while rapidly recovering to a desirable operational state (recovery), including complex coordination between entities and operation when information on the nature of an attack is incomplete.

Incorporating specification and verification techniques into design cycles to bring the benefits of verification approaches to large-scale infrastructure systems is needed.

Broadly speaking, measurement and verification operations are conducted to obtain answers to particular questions and/or assist with the decision-making processes. In the context of security of CPHS infrastructure, new measurement tools and metrics can help determine whether a system has reached a bad state (e.g., violation of well-defined safety specification), is freed from a specific class of vulnerabilities (e.g., by monitoring system response after deployment or update of defense tools), or is able to identify that it has been tampered with (e.g., by tracking and refining adversarial signatures). To make progress on these issues, we need both theoretical advances and computational tools to support the design of provably secure and dependable systems.

Although much progress has been made in the past decade on new ways to measure, evaluate, and verify security, challenges still remain because of the distinctive features of CPHS. These challenges arise because CPHS operate in highly dynamic and uncertain environments and rely on complex hardware-software interfaces supporting multi-scale interactions between humans, AI algorithms, hardware, and networks. Continuous monitoring and principled verification tools of these interfaces in the face of the changing threat landscape and unpredictability of overall system behavior is, therefore, one of the outstanding challenges.

Establishing new metrics that build on measurement and verification efforts can help address both technological and institutional bottlenecks for security solution deployment. Consider the question: for well-defined threat model(s), *how much security is desired* and *how many resources (budget, human effort) need to be allocated to achieve a certain improvement, as in an appropriate security metric?* To answer this question, a principled set of tools is needed to identify where to deploy measurement technologies and which quantities to measure based on the applicable notions required for security and trust. Additionally, most system architectures would require derived measures based on how components interact with each other and the nature of uncertainty propagation in response to adversarial models, both independent and correlated. Furthermore, we can also expect that human-centric systems involving AI or machine learning (ML) algorithms and automated processes will require new ways to model and predict both individual actions and collective behavior.

The visioning event participants defined the need for security metrics that enable quantification of tradeoffs between usability, security, and return on investment. That is, future research on metrics should examine how to balance the usability of a system with the cost of implementing security measures to protect it. Furthermore, it is important to determine whether such measures alert external agents to the presence of new vulnerabilities or enable the system to self-monitor and respond to issues when they arise. A continuing issue with metrics remains whether it is possible to create consistent units of measurement – often the aforementioned tradeoffs are hard to evaluate due to heterogeneities in ways of evaluating usability and security. Of course, adoption of new (and hopefully more effective) metrics is contingent on a clear demonstration of when, why, and how conventional metrics fail.

Three topics have emerged that could have the greatest return on investment and should be prioritized by the engineering research community.

Infrastructure designs to make problems pop: Traditional control system design principles of observability and controllability can provide a conceptual approach to integrate observability of the security state of the CPHS as a first-class design requirement. Observability in the context of adaptive cyber defense is an inherently challenging problem. The defenders typically have incomplete information about the type of adversaries and need to consider the adaptive nature of attacks (i.e., considering that sophisticated adversaries will learn from and attempt to evade defenses).

Foundational research and practical tools are thus needed to observe, estimate, and update the dynamic security state of a CPHS and use state information for improved situational awareness, cyber deception, and response to detected malicious behaviors. Measurable progress to address this challenge can bring significant benefits, as *complex networks are becoming even more prone to autonomous adversaries*.

Models for fully autonomous attack recovery: Future CPHS will face a multitude of attacks and failures. Researchers must design fully automatic mechanisms that can maintain key functionality (resilience) while rapidly recovering to a desirable operational state (recovery). There is a need to develop practically effective and tractable ways to facilitate optimal recovery rates after disruptions, recognizing that critical infrastructures under attack must be protected during narrow windows of opportunity to prevent local failures from escalating into global (network-wide) disruptions. Furthermore, recovery operations often need coordination between multiple entities, given their individual objectives and imperfect information about the nature of attacks.

A potential avenue for future research is how to integrate predictive information about attacks (and their potential impacts) into autonomous decision-making agents that can trigger additional information-gathering tasks (e.g., through activating sensors) and defense mechanisms to reduce time-to-recovery and prevent system-wide failures.

Incorporating specification and verification techniques into design: A system without a security specification cannot be insecure; it can only be surprising. Infrastructure systems must be trustworthy, and trustworthiness arises from predictability. Today, systems frequently lack comprehensive specifications, in part because of the difficulty of writing those specifications and keeping them in sync with system evolution. Evidence is available now for the effectiveness of specification and verification techniques in defeating attacks, but the techniques must be substantially advanced and incorporated into normal design cycles to bring the benefits to large-scale systems. This effort could be combined with a related one in architecture that focuses on segregating critical (to be specified and verified) parts of systems from non-critical parts.

Broadly, research is needed in incorporating specification and verification techniques into design cycles, advancing the methods used, and increasing the scope of system properties within CPHS infrastructures.

#3

Future Approaches to Autonomous Security

SUMMARY OF RESEARCH PRIORITIES

Autonomous security in CPHS infrastructure is needed to address both the scale of tomorrow's system infrastructures and the complexity of adversarial threats.

Research should include how intelligent automation and human intelligence interact, and how the appropriate synergistic balance can make security more powerful, adaptive, and intelligent.

The future of AI-driven security research is to add automated decisions and response to today's statistical modeling and predictive approaches.

A key challenge in future autonomous security is the need for more sophisticated contextual awareness. Tomorrow's autonomous security should have a rich understanding of devices, infrastructure context, operator intentions, and more to make or recommend almost human-like decisions.

Key applications for tomorrow's autonomous security include virtual security assistants, automated configuration agents, real-time security risk analyzers, and adversarial agents for use in design analysis and vetting.

Another major area for future investment in engineering research and development for unhackable infrastructure is *autonomous security*. Here, autonomous refers to technologies that are self-configuring, self-guiding, self-managing, self-learning, self-tuning, and so on. Modern examples of autonomous systems include industrial robots, self-driving vehicles, and modern avionics. While some technical domains have a long history of autonomous systems research, we believe that applications to infrastructure security are underexplored.

There are many reasons why autonomous systems are sorely needed to enable security in tomorrow's cyber-physical-human infrastructures. The first is *scale*. As countless physical domains become more fully digitalized (i.e., cyber-physical), the complexity and scale of cyber-controlled (and cyber-attacked) infrastructure become immense. Human monitoring, threat detection, and threat response will simply become infeasible at scale without the help of autonomous systems.

A second reason is *threat complexity*. As adversaries become more sophisticated in the tools they use and more strategic in their approaches, threats become less observable by human operators, and the need for well-constructed autonomous detection systems increases. Autonomous systems, for example, will leverage extensive data collection and monitoring capabilities and combine them with sophisticated analysis techniques. Examples of complex threats include multi-phase attacks, distributed attacks with no single adversarial source, and multi-attacks that obfuscate the features of any single component attack.

A third reason for autonomous security is *to enable a more powerful human response*. While automation and autonomous systems are commonly understood to mean the replacement of humans with artificial, engineered substitutes, the two are more likely to be used synergistically. That is, autonomous security may automate complex or tedious processes and functions that computers do well in the service of humans. For example, automated agents can be deployed to collect and aggregate data about the infrastructure, analyze complex data statistically, perform detailed verification tasks, or identify anomalies *so as to inform* human operators of when, where, and at what scale a threat has been observed. Autonomous subsystems may present alternatives for human decision-making.

Human Synergies with Autonomous Security

There is a need for far richer research literature on how automation and human intelligence interact within security, and how the right synergistic balance could make future infrastructure security more powerful, adaptive, and intelligent. There is a need to determine which domains can be partially or fully automated and which cannot, as well as which characteristics of the human element are irreplaceable and ways to understand them within CPHS design. Systems may be partially autonomous but still rely on human intention, human decision-making, and human analysis—for example, giving the human operator oversight when a false positive or negative has occurred. A human operator may need to override autonomous security actions when factors are beyond the design specifications of the automation.

A much-needed interactive technology for human infrastructure designers is a *design-time security risk analyzer*. Instead of building an infrastructure and then identifying weaknesses after deployment, researchers could develop an autonomous security technology that allows a human designer to interact iteratively by inputting proposed design specifications and receiving as output an analysis of security risks. Security could be analyzed at multiple levels, from detailed component or interface designs to the potential for vulnerabilities in larger subsystems, and to the system and its design as a whole. Risks could be categorized by severity or type, and specialized taxonomies (e.g., the MITRE ATT&CK framework[3]) could be developed to inform the analysis.

The above technology can be generalized into a larger research challenge to build dynamic virtual security assistants. Such an autonomous security agent could help humans within CPHS infrastructures to better observe and reason about security. The tool could have both natural language and visual user interface components to convey rich information sets, and perhaps include immersive AR/VR modes that are highly adaptive to the tool's human user. The idea could be applied in a range of scenarios, from naive users who need coaching and reference information to security operations staff who ask sophisticated questions about the state of the system, perform forensic investigations, or plan infrastructure modifications. The assistant would dynamically learn over time and increase its expertise in infrastructure monitoring and analyzing tasks for human consumption.

Finally, research is needed on autonomous security that takes human intentions and translates them to infrastructure actions. As cyberinfrastructures become both increasingly scaled and complex, autonomous security could help to navigate the extensive detail involved in configuring security controls, monitoring across heterogeneous hardware domains, and analyzing the meaning of low-level system information (e.g., logs and error messages). Autonomous security tools could help implement low-level design given high-level human specifications or implement domain-specific programming for low-level devices.

AI Challenges in Autonomous Security

Workshop participants discussed the overall capability progression of today's AI-enabled systems and what tomorrow's systems could look like. Today, there is an explosion of applications leveraging ML techniques to make systems more predictive and anomaly-aware. But the future is for researchers to add automated decisions and responses to the statistical modeling to create truly autonomous systems. In security, autonomous systems could provide real-time defenses against a newly discovered threat or modify moving target algorithms automatically to counter a pattern in adversarial behavior.

A key research challenge is greatly expanded contextual awareness. Modern infrastructure devices perform various functions but seldom know much beyond narrowly defined input sources and sensors. There is a pressing need to develop ways for systems to comprehend real-world device contexts more fully and to make more insightful, automated decisions with respect to security and resilience. For instance, a system should be more aware of operator actions and capable of distinguishing them from adversarial actions. Ideally, future autonomous security would have a rich understanding of infrastructure components and context and make or recommend almost human-like decisions. Extensive research is also needed for improving the level of contextual understanding in automated risk analysis.

Another research challenge for tomorrow's AI-based autonomous security is leveraging multi-modal data sources. Today's ML-based security is often focused on model training with single data sets from a single device type or source. Future security should ingest data from across the entire infrastructure and leverage new data science techniques to model adversarial behavior across a more comprehensive and multi-modal set of data sources. Filtering noisy data for salience, managing models at multiple levels and time granularities, and enabling analysis to run in real time are all key challenges in this vision.

The ultimate goal in AI-driven autonomous security systems is human-like decisions and responsiveness. Humans are endlessly adaptive, evolving in their capabilities, circumspect (at their best) and often non-binary in their decision-making. A key challenge for future autonomous security is human patterns of self-learning. Faced with a new technology domain, humans recognize when it is time to get more information and extend their understanding before making a decision or taking an action. Future autonomous security should do the same.

Anticipating Adversarial Use of Autonomous Systems

Every discussion of security technology advancement must consider the parallel advancement in adversarial technology and tools. As such, a key area of research on autonomous security should be exploration of adversarial usage and the implied defenses. For example, fully autonomous systems create the possibility of new, more powerful malware and "bot" technologies, or the application of AI to learn and evade threat detection mechanisms.

A potential line of research is the **use of adversarial autonomous systems for design analysis and vetting**. As researchers investigate adversarial technology use, they could apply it within tools that test the security robustness of a proposed infrastructure design. Conversely, designs could be used to explore the limits of adversarial capabilities to learn rule-breaking patterns.

Autonomous Security Architectures

Future research on autonomous security should consider **new architectures and design principles that rethink how security can be directly designed into future infrastructures**.

One example is what discussants referred to as a *hierarchical, dynamic component manager (HDCM)* framework. The idea is for an autonomous agent to manage activation and availability of system components on an as-needed basis. Instead of building systems that offer hundreds or thousands of capabilities and accompanying APIs, an HDCM framework could select from archived components only those that are minimally needed to carry out the needed modes of functionality. Using this strategy, systems could minimize their attack surface and reduce the complexity of security monitoring and threat response. Minimal use architectures could be a future design pattern.[4]

#4

New Approaches to Resilience in Interdependent Infrastructures

SUMMARY OF RESEARCH PRIORITIES

A key design challenge is managing insecurities arising from correlated software bugs and hardware malfunctions. CPHS tightly couple continuous physical dynamics with networked computer processes. Adversaries can exploit the link between reliability and security.

Research is needed on the complex interplay between coordinating entities in CPHS infrastructures. Large-scale infrastructures require coordination and compliance across profit-driven entities, and research is needed on interdependency risks, architectures, and operational tools.

An important opportunity for future research is to develop a design approach that maintains system-level properties of safety and security after integration of modular components

Compositional and learning-based approaches are needed to measure and quantify system-level safety properties based on data-driven and stochastic models of CPHS.

Tomorrow's systems will be deployed in contested environments that require far more active cyber defense strategies and tactics. This includes resilience to noisy and untrusted data, and adaptive response to evolving adversarial capabilities.

Today's critical infrastructures are spatially distributed across large physical areas and consist of heterogeneous cyber-physical components interconnected by communication networks with complex peering and hierarchies. In the past two decades, significant progress has been made on improving the cyber-security posture of classical supervisory control and data acquisition (SCADA) systems and modern networked control systems deployed in industrial facilities. Yet significant challenges remain in assessing and managing security risks to CPHS and developing proactive mechanisms to improve their resilience to correlated disruptions, both random and adversarial.

The first challenge arises from the complex nature of threats driven by the exposure of infrastructures to insecurities from correlated software bugs and hardware malfunctions. Attackers (both external hackers and malicious insiders) are becoming more adept at exploiting the vulnerabilities of hierarchically managed CPHS and compromising critical infrastructure operations. These issues are exacerbated by the heterogeneity of and often lack of enforcement mechanisms for access control and authorization among the strategic players such as the operators, the SCADA and security vendors, and the end users of the system.

The danger of correlated failures becomes especially profound in CPHS due to the tight coupling of continuous physical dynamics and discrete dynamics of embedded and networked computing processes. Such failures can also increase the risk of cascading failures (e.g., power network blackouts affecting communication networks, and vice versa), with huge societal costs. Furthermore, in these situations, it is extremely difficult and costly to isolate the cause of any specific failure using the diagnostic information which, in general, is imperfect and incomplete. Hence, reliability and security failures in CPHS are inherently intertwined, and major research progress needs to be made to analyze and manage the risks of interdependent reliability and security failures.

A second challenge results from the complex interplay of technological defenses and incentive structures in all modern infrastructure domains. Large-scale critical infrastructures are managed by profit-driven, private entities. It follows that the implementation of technological defenses and resilient system-wide operations require compliance and coordination among these entities. In recent years, several industry and government organizations have developed security standards and recommendations that combine security tools (e.g., access control mechanisms, network intrusion detection systems, and security certification) and control systems tools (e.g., model-based attack diagnostics, robust and fault-tolerant control, and adaptive/reconfigurable operations). However, in practice, the information about system failures as well as incentives to ensure security vary significantly across different entities.

Interdependent failures in CPHS can result in misaligned player incentives, i.e., the individually optimal security defenses diverge from the socially optimal ones. In fact, even in the presence of advanced technological defenses and institutional measures, the residual risk due to correlated failures cannot be completely eliminated, thus requiring new ways to benchmark these risks, reshaping the functional architecture of infrastructures, and developing tools for resilient and adaptive operations.

Design of Resilient Infrastructure Systems

The designers of modern infrastructure systems approach the monitoring, control, and management tasks using a layered architecture comprised of regulatory control, supervisory control, and management level. This architecture is well suited for composition of multi-level controllers and permits defense schemes based on hardening of individual components as well as securing communications between them. **An important opportunity for future research is to develop a design approach that maintains system-level properties of safety and security after integration of modular components.**

It is important to distinguish between *robustness* – which requires maintaining a certain level of performance in the face of a class of perturbations, and *resilience* – which requires timely and often dynamic (multi-stage) response to perturbations in an effort to limit the magnitude and/or duration of reduced performance. New approaches to maintain resilience in future infrastructures must account for the above-mentioned interdependencies and informational deficiencies. A comprehensive design approach must be well-integrated with AI-driven tools for detection and identification of attacks/faults, safety-preserving response mechanisms, and learning and adaptation based on incremental and noisy information.

Measuring and Quantification of Resilience and Risk

A significant opportunity for innovation is a quantitative risk and resiliency assessment framework based on data-driven and stochastic models of CPHS that accounts for interdependent reliability and security failures and yet does not require costly explorations of the enormous design space. Indeed, component-wise sensitivity analysis of how the system responds to perturbations can quickly become intractable even for small-scale systems. Evaluating and verifying safety property – defined as the ability to maintain the state of CPHS within a set of desirable states – in the context of infrastructures with varying degrees of human-AI interactions is also an important area of research.

Such an agenda may require a compositional and learning-based approach to measure and quantify system-level properties. To rigorously evaluate these properties, quantitative tools from statistical estimation, model-based diagnostics, stochastic simulation, and adaptive control are relevant. These techniques can also allow synthesis of effective security strategies and provide estimates of failure probabilities in highly uncertain environments that future infrastructures will face.

Active Cyber Defense and Adaptive Response Operations

Our infrastructure systems are steadily becoming more autonomous to support high-demand capabilities such as driverless transportation, flexible energy systems with mobile generation systems, and autonomy-driven emergency operations. The scale, scope, and pace of autonomous systems are steadily outperforming humans and human-controlled systems. These capabilities are also making attacks more advanced, persistent, adaptive, and even deceptive. Thus, future autonomous systems deployed in highly contested environments will require new cyber defense strategies and tactics, and associated tools and procedures.

In particular, we need new tools for learning and adaptation that can provide advanced capabilities such as reasoning about situation awareness and dynamic alert prioritization based on noisy and untrusted data, cyber deception to shape the information structure of adversaries and potentially induce them to take erroneous and less impactful actions, and adaptive response operations that account for evolving adversarial capabilities.

This agenda will need to go beyond traditional cyber defense in information systems and consider the emerging cyber-physical network architectures in mixed-autonomous systems. Additionally, innovations that support cyber diversity, redundancy, and randomization are needed to enable dynamic and adaptive manifestations of CPHS, creating new ways to protect and defend these critical systems and making it harder for the attackers to learn, reverse engineer, and exploit the system in comparison to a fixed target.

Credit: Mathis Mousavi, Unsplash

#5

Architecting Trustworthy Systems

SUMMARY OF RESEARCH PRIORITIES

Transforming ill-defined notions of trustworthiness (e.g., vendor reputation) into more well-defined, robust notions of proven correctness and security.

Design specifications that play a central role in verifiable CPHS. Issues include specifying the intractable space of possible incorrect system actions, specifying threat models more rigorously, and identifying adversarial abuses that remain within system specs.

Research principles about the nature of security and reliability in both centralized and decentralized infrastructure contexts. A key issue in tomorrow's CPHS infrastructure is centralized versus decentralized control.

Scaling confidential computing techniques (e.g., attestation, isolation) defined in single system contexts to larger infrastructures with complex hierarchies of components and cross-domain interactions.

Trustworthy architectures for a large number of new infrastructure domains. Some of these include AI-based agents, human augmentation (clothing, physical implants, AR/VR), autonomous transportation, and infrastructure (water, food, power, other utilities).

Applying quantum-resistant cryptography to future CPHS infrastructure and determining where it is needed and how systems will be impacted by the required resources to develop and implement it, especially in small devices and embedded systems.

A key prerequisite for unhackable infrastructures, and a major area for future investment in engineering research and development, is *trustworthy system architectures*.

The notion of "trust" and "trustworthy" requires some explanation. Trust is commonly thought of in distinctly human terms. People *trust* select people or sources whom they believe are reliable, truthful, and sincere. People who have these characteristics are seen as *trustworthy*; they are deserving of trust because their conduct and attributes warrant it. With trust in place, we can do business together and collaborate on private matters.

In the context of unhackable systems, *trustworthy* refers to system correctness and reliability. A *trustworthy system* is one that functions exactly as expected within its context of use, even when unexpected inputs or failures occur. There is a distinction between *trustworthy* and *secure*. A *secure system* is one that continues to be trustworthy while under attack. That is, it maintains its correctness and reliability even as an adversary attempts to gain entry, manipulate an interface, execute malware, or hijack system control.

Research in trustworthy systems and architectures examines how correctness and resilience can be built into a system architecture, and how trustworthiness can be proven. The challenge is to transform vague notions of human trust (e.g., we have confidence in a vendor, the system appears to be functioning normally) to more rigorous notions of operational health and correctness — especially freedom from adversarial compromise. Illustrative research questions include: How do we know a system is functioning correctly? What verification methodologies can be devised to prove operational correctness given the system state? How do we know a system has not been compromised? How should a system be built to better enable monitoring and robust evaluations of correctness?

Design Specification

Notions of correctness and security within trustworthy infrastructure imply the central role of a design specification. To understand whether CPHS function correctly, their behavior must be compared against a formal description of

legal system states, and what the system is supposed and not supposed to do. A key area of research investment should thus be on defining correct behavior in complex infrastructures with many component sub-infrastructures, and in devising verification methodologies that comprehend this complex array at all required levels (individual components, component interactions, high-level system function, layers of cross-system functionality, etc.).

Defining what a system shouldn't do is a difficult challenge. While the space of correct system actions may be large, the space of possible incorrect actions is much larger and seemingly intractable. Yet this is precisely the space adversaries occupy as they look for unexpected vulnerabilities and abuses. The goal is to design and build trustworthy systems that are resilient to innumerable malicious inputs and crafted modes of system abuse intended to derail system function and/or hijack system control. **The challenge is how researchers can better address the problem of specifying system constraints to manage the vast space of inputs and manipulations that are simply out of spec within CPHS infrastructures.**

A closely related problem is that of specifying threat models within complex CPHS infrastructures. Security risks and threats are often considered in terms of malicious agents and adversarial objectives. But these high-level notions need more specific and formal definitions that can be applied to the context of design and verification. For all the attention security receives as the paramount consideration in system design and deployment, it is surprising that *what a system is protecting against* is often so poorly discussed and underspecified. Research is needed in CPHS infrastructures to create more robustly defined threat models.

Two research areas noted by MIT's David Clark include devising mechanisms that distinguish different classes of system use and isolating impactful design components. As previously noted, many adversarial actions represent unanticipated uses of a technology that are entirely *within* a system's functional specifications. To better address this problem, tomorrow's systems could define fine-grained classes of system uses and identify human intention in ways that don't simply presume a trusted operator. To address the problem of infrastructure complexity, tomorrow's CPHS infrastructures could be designed to isolate aspects of a system that have the most potential to cause harm and to design-in architectural features facilitating monitoring and verification.

Centralized Versus Decentralized Control

One aspect of CPHS infrastructure trustworthiness that needs attention within engineering R&D is centralized versus decentralized control. On the one hand, decentralized control helps to mitigate the problem of vulnerabilities leading to infrastructure-wide compromise. An adversary who successfully exploits a centralized control vulnerability may gain access to and control over the entire CPHS infrastructure. Centralized control also provides opportunities for insider threats and magnifies their impacts. Decentralized control can often provide more checks and balances, e.g., consensus protocols that prevent a single rogue agent from sabotaging a collective system action.

On the other hand, decentralized systems often evolve slowly because of the need for consensus among participants. They may lack the right unifying abstractions and control infrastructure, and associated security protections. Another problem is the potential for adversarial design participants who have malicious intentions unbeknownst to other members. Operationally, a system may be less efficient given the need for extensive "zero trust," "trust but verify," or group consensus mechanisms.

In the end, however, it may be that centralized control is a necessary requirement for effective threat detection, mitigation, and response. **Research is needed to reason in a more principled manner about the nature of security and reliability in both centralized and decentralized CPHS infrastructure contexts.** Questions for discovery include whether both are needed but at different levels; whether open architectures can provide unifying frameworks for decentralized design; and whether a more fundamental analysis of distributed systems, along with well-defined adversarial models, can shed light on this issue.

Confidential Computing Approaches to Trustworthiness

Another area for engineering research is applying confidential computing principles and technologies to CPHS infrastructures. Confidential computing techniques use hardware platform features to provide physical isolation from an adversary on the same system. For example, a trusted execution environment may offer isolated memory partitions and central processing unit features that protect control software from being exposed to hackers running malware on the same system.

A key feature of confidential computing is *attestation*. Attestation is a robust, platform-based capability in which hardware mechanisms hash and digitally sign software, data, and/or configuration running on the system. Using cryptographic techniques (e.g., public key cryptography or secure communications protocols), platform-derived "proofs" provide evidence that system software has not been compromised and that the systems state remains within expected parameters.

Trustworthy CPHS infrastructures of the future should apply confidential computing techniques extensively to build robust verifications for system components, and for raising the bar on verification between interacting components. But research is needed to transform techniques defined in single system contexts to larger infrastructures with complex hierarchies of components and cross-domain interactions. Research questions may ask how such infrastructures can be instrumented to perform confidential computing verifications across heterogeneous hardware platforms; how rigorous yet performant verification systems should be constructed; and how verification failures should be handled without terminating system operation.

New Infrastructure Domains

Finally, future research on trustworthy CPHS infrastructures will need to address many new domains. Expanding on recent work in explainability, approaches are needed for making AI agents within tomorrow's CPHS infrastructures trustworthy. Tomorrow's AI agents will be increasingly sophisticated in scope, data sources, decision domains, and more. Questions remain concerning how to verify their function in robust ways and the designs that would make them amenable to verification.

As humans become more deeply integrated into CPHS infrastructure, they will use technology to augment clothing and even the physical body. For example, humans may rely on medical implants or wearables that manage physical function or alleviate discomfort. Or they may use both augmented and virtual reality to interact with technology or their social "metaverse." Research is needed to develop devices and modes of interaction that are verifiably trustworthy, even under attack.

Tomorrow's infrastructure will be increasingly autonomous. For example, autonomous flight and automotive transportation will become routine. As CPHS infrastructures become increasingly autonomous, there is an urgent need to develop approaches and technologies that ensure trustworthiness. Systems will be complex in their functionality and operate in real time. Once again, there is a need to ensure infrastructures remain reliable in the face of both cyberattacks and physical attacks (e.g., analog sensors, elements of feedback control systems).

Trustworthiness safeguards and techniques are also needed for infrastructures underpinning clean water, food engineering, utilities (e.g., electrical grids), and the environment. Interestingly, adversarial elements in these domains can be either human or nature itself. The latter is capable of causing catastrophic harm through severe weather, biological outbreaks, resource scarcity, and more.

The emergence of scaled quantum computing, while an exciting technology for scientific computing, represents a threat to the trustworthiness of our current cryptographic protections. Research is needed on applying quantum-resistant cryptography to future CPHS infrastructure and to determine where it is needed and how systems will be impacted by the required resources to develop and implement it, especially in small devices and embedded systems.

Overall Assessment and Moving Forward

Each focus area discussed at the workshop provides opportunities to progress toward unhackable infrastructure through proactive and comprehensive research approaches. Encouraging the engineering research community to pursue research directions for cybersecurity provides an opportunity for federal agencies and the entire engineering research community to invest their resources and talent to address this problem with broader impacts on society.

Three high-level themes were selected as broad research principles by participants:
- integration of security into design cycles;
- infrastructure transparency; and
- improving the human condition through the lens of unhackable infrastructure.

The urgent need for early, proactive **integration of security into design cycles** is present in every research priority—the imperative that solutions in cybersecurity can no longer simply react to yesterday's crisis. Participants underscored the need to look forward and to integrate security into design cycles of systems—baked-in rather than bolted-on. For example, autonomous systems must be designed to include the necessary contextual awareness and data sources for threat detection and response. Observability must be intrinsic to infrastructure design; administrators of infrastructure systems must always understand system status, and adversaries must never be able to camouflage system status.

Infrastructure transparency takes many forms and represents a second core goal in cybersecurity research and development. The ability to measure the state of a system and to make problems obvious if a system has been attacked or otherwise compromised must be built in. A transparent system can also be a more trustworthy system; too often, security is assumed because no indications arise to the contrary. But systems without transparency will only surprise—and not always in beneficial ways. All who rely on key infrastructure systems as users or administrators must know that such systems are trustworthy, and transparency is key to ensuring this.

The third core goal for cybersecurity research is that of **improving the human condition through the lens of unhackable infrastructure**. Whether transportation, essential utilities, or medical devices, security should enhance the trustworthiness and reliability of devices and infrastructure serving humans. Developing secure and effective medical implants or wearables is one obvious technical domain, as are personal security assistants to help users across a spectrum of knowledge levels make informed security decisions.

Anticipating and preparing for security threats, whether short-, medium-, or long-term, has escalated as a critical national and international priority. Engineering-led research and innovation in these priority areas is essential to lay the foundation for the high-impact, multidisciplinary research necessary to mitigate the economic and security threats inherent in cybercrime and secure both physical and virtual spaces.

Appendix A: Visioning Event Participants

Saurabh Amin, MIT

Gregory Buck, Gray Space Strategies

Alvaro Cardenas, University of California- Santa Cruz

Kenneth Carnes, Tennessee Valley Authority

David Clark, MIT

Mathieu Dahan, Georgia Institute of Technology

Lynn Davis, Research Triangle Institute

Kevin Fu, University of Michigan

Luis Garcia, University of Southern California Information Sciences Institute

Sean Guillory, Booz Allen Hamilton

Jennie S. Hwang, H-Technologies Group

Todd Jones, Sandia National Laboratory

Bashir Khoda, University of Maine

Carl Landwehr, University of Michigan

Hua Harry Li, San Jose State University

Vinton Morris, Morgan State University

Hieu Nguyen, North Carolina A&T State University

David Ott, VMware

Ashley Podhradsky, Dakota State University

Zhihua Qu, University of Central Florida

Priyank Srivastava, MIT

Vipin Swarup, MITRE Corporation

Shambhu Upadhyaya, University at Buffalo

Urjaswala Vora, Penn State University

Weichao Wang, University of North Carolina- Charlotte

Spencer Wilcox, NextEra Energy

Zhenhua Wu, Virginia State University

National Science Foundation Observer

Louise R. Howe

Contractual Partners

Carolyn Kaldon, University of Memphis, Center for Research in Education Policy

Todd Zoblotsky, University of Memphis, Center for Research in Education Policy

Susan Lang, The Ohio State University

ERVA Principal Investigators and Operations Team

Tony Boccanfuso, UIDP

Charles Johnson-Bey, Booz Allen Hamilton

Mark McGill, ERVA

Annie Shealy, ERVA

Kristina Thorsell, UIDP

Appendix B: Event Presentation Summaries

Panel: Engineering R&D Solutions for Unhackable Infrastructure

Moderated by Saurabh Amin, MIT, with panelists David Clark, MIT; David Ott, VMware; and Vipin Swarup, the MITRE Corporation

This panel framed subsequent breakout session conversations, providing each of the three panelists an opportunity to raise key issues facing research and development in infrastructure cybersecurity. Discussion topics fell under three general areas: **the dynamic, uncertain, highly strategic future of cybersecurity; design specification/system architecture; and risk assessment and mitigation.**

The Dynamic, Uncertain, Highly Strategic Future of Cybersecurity

Decades of work on insecurity and resilience haven't resulted in hoped-for gains, in part because security and resilience for infrastructures are moving targets. All components—software, devices, contextual elements, users, and adversaries—adapt and evolve. Time is a parameter of both security and resiliency; adaptive, scalable security requires that systems are built with the understanding that some incursions will succeed. But systems must also build in resilience and recovery because, by the time a system is configured to respond to a threat, the threat has evolved. There is a need to better anticipate future threat landscapes instead. Examples include designing systems that better insulate users from threats, building more agile systems for monitoring and response, or creating more programmable security rather than simply relying on configurable security.

Design Specification/System Architecture

How usability is or should be an element of security must be considered. Security suffers from a combination of legacy and future drag. Legacy drag is caused by a plethora of flawed devices and systems that are still in use and litter the infrastructure landscape. They pose the real-world challenge of how to deal with legacy technology, and how to manage backward compatibility with poor design choices. Future drag refers to how advances in security are slowed by a lack of economic incentives to invest. Security must be recast as a quantifiable feature possessing intrinsic value ripe for investment—not just necessary fixes or ad hoc responses to newly discovered system vulnerabilities. Ultimately, building in security and resilience is a combination of top-down and bottom-up approaches. Principled approaches should be developed that guide design in top-down and reusable ways. At the same time, each system poses specific architectural challenges and requires bottom-up, system-specific design reasoning. Trustworthiness often requires low-level features, which provide building blocks for high-level verification systems. Low-level security operational details are also needed by decision makers at a high level. Finally, whether secure systems can be created with decentralized control remains unanswered.

Risk Assessment/Mitigation

In a narrow sense, when one defines a secure system, one refers to a system that does what it should do, even when under attack. One might say that a system built without specifications can never fail because no one has quantified what it is supposed to do. It's critical to understand what a system is designed *to* do, designed *not* to do, and what it *can* do that is potentially harmful (within specific parameters). Disinformation can create undesirable behaviors even when using a system as specified.

Another obstacle to creating scalable security is procuring data from private entities and systems to enable research. If goals for more effective risk assessment and mitigation are to be met, it is critical to understand system dependencies and how many roles and agents converge in modern infrastructure.

References

1 Center for Strategic and International Studies. (2022, July). Significant cyber incidents. Significant Cyber Incidents. Retrieved September 1, 2022, from csis.org/programs/strategic-technologies-program/significant-cyber-incidents

2 Morgan, S. (2021, April 27). Cybercrime to cost the world $10.5 trillion annually by 2025. Cybercrime Magazine. Retrieved October 6, 2022, from cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/

3 MITRE ATT&CK® is a globally-accessible knowledge base of adversary tactics and techniques based on real-world observation. See attack.mitre.org

4 Design patterns are "descriptions of communicating objects and classes that are customized to solve a general design problem in a particular context." Gamma, E., Helm, R., Johnson, R., Vlissides, J. (1998). Design patterns elements of Reusable Object Oriented Software. Addison Wesley

erVa

NSF Engineering Research
Visioning Alliance

Our mission is to identify and develop bold and transformative new engineering research directions and to catalyze the engineering community's pursuit of innovative, high-impact research that benefits society.

ERVA IS FUNDED BY THE NATIONAL SCIENCE FOUNDATION THROUGH AWARD NUMBER 2048419

This material is based upon work supported by the National Science Foundation under Grant # 2048419. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Ervacommunity.org