Distribution Statement A: Approved for public release; distribution is unlimited.

VOLUME 8 | NUMBER 2 | 2024

CSIAC JOURNAL

PAGE 24

NAVIGATING CHALLENGES AND OPPORTUNITIES IN THE CYBER DOMAIN WITH

SIM2REAL TECHNIQUES

DEPARTMENT OF DEFENSE
UNITED STATES OF AMERICA

CSIAC JOURNAL

Volume 8 // Number 2 // 2024

Editor-in-Chief:
Aaron Hodges

Sr. Technical Editor:
Maria Brady

Graphic Designers:
Melissa Gestido, Katie Ogorzalek

The CSIAC Journal is a publication of the Cybersecurity & Information Systems Information Analysis Center (CSIAC). CSIAC is a DoD Information Analysis Center (IAC) sponsored by the Defense Technical Information Center (DTIC) with policy oversight provided by the Office of the Under Secretary of Defense (OUSD) for Research and Engineering (R&E). CSIAC is operated by the SURVICE Engineering Company.

Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or CSIAC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or CSIAC and shall not be used for advertising or product endorsement purposes.

ISSN 2836-7383 (Print) // ISSN 2836-7391 (Online)

Distribution Statement A:
Approved for public release; distribution is unlimited.

On the Cover:
Digital Art Rendering *(Source: zabelin [123rf.com])*.

CSIAC Cybersecurity & Information Systems Information Analysis Center

ABOUT CSIAC

Who We Are

A DoD Information Analysis Center comprised of scientists, engineers, researchers, analysts, and information specialists.

What We Do

Generate, collect, research, analyze, synthesize, and disseminate scientific and technical information (STI) to DoD and federal government users and industry contractors.

Why Our Services

To eliminate redundancy, foster collaboration, and stimulate innovation.

CSIAC SERVICES

Subject Matter Expert (SME) Connections

Access to a network of experts with expertise across our technical focus areas.

Technical Inquiries (TIs)

Up to 4 hours of FREE research using vast DoD information resources and our extensive network of SMEs.

Specialized Task Orders

Research and analysis services to solve our customer's toughest scientific and technical problems.

Webinars & Events

Our webinars feature a technical presentation from a SME in one of our focus areas. We also offer key technical conferences and forums for the science and technology community.

STI Collection

Our knowledge management team collects and uploads all pertinent STI into DTIC's Research & Engineering Gateway.

Information Research Products

The Cybersecurity & Information Systems Digest, state-of-the-art reports, journals, TI response reports, and more available on our website.

CONTACT CSIAC

IAC Program Management Office

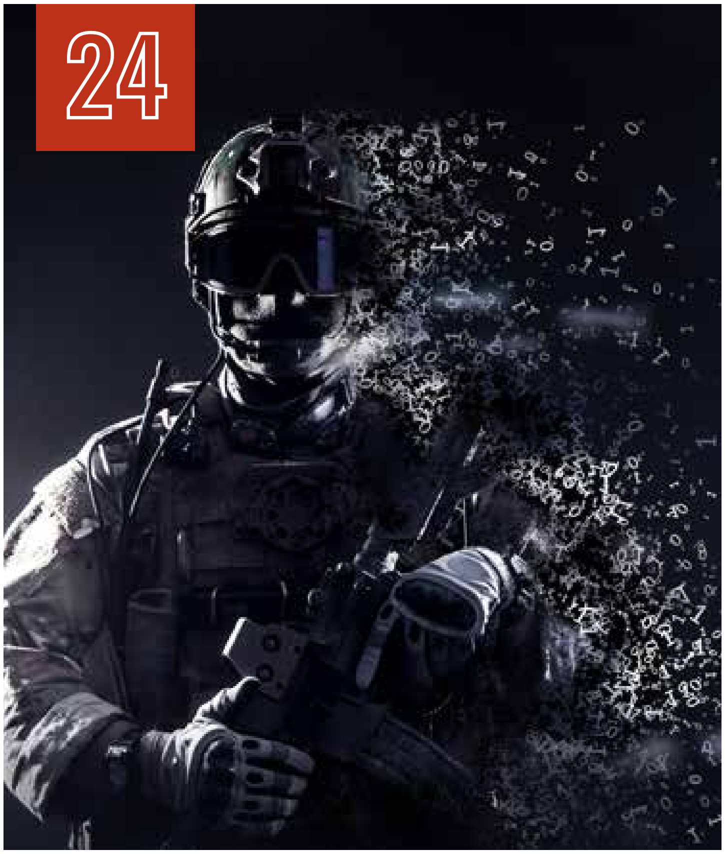
8725 John J. Kingman Road
Fort Belvoir, VA 22060
Office: 571.448.9753

CSIAC Headquarters

4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600
Fax: 410.272.6763
Email: contact@csiac.mil

CSIAC Technical Project Lead

Phil Payne
4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600



FEATURED ARTICLE

NAVIGATING CHALLENGES AND OPPORTUNITIES IN THE CYBER DOMAIN WITH SIM2REAL TECHNIQUES

Emily A. Nack and Nathaniel D. Bastian

This article examines the emerging intersection of Sim2Real techniques and the cyber realm, exploring their challenges, potential applications, and significance in enhancing our understanding of this complex landscape.

IN THIS ISSUE

EXTREME ZERO TRUST

BY DAKSHA BHASKER AND CHRISTOPHER ZARCONE

(PHOTO SOURCE: DROGATNEV [CANVA])

INTRODUCTION

Zero Trust Architecture (ZTA) has become a mainstream information security philosophy. Many commercial enterprises are in varying stages of their journeys in adopting and implementing ZTA. Similarly, federal policy has moved toward ZTA, motivated by actions such as Executive Order 14028 and guided by NIST 800-207 and CISA's Zero Trust Maturity model [1–3]. Although there are many frameworks for ZTA in industry, academia, and government, the essence of the philosophy can be summarized in a few brief maxims (Figure 1), such as the following:

- Never Trust, Always Verify
- Least Privilege Access
- Assume Breach

ZTA has spawned architectures and implementations that vary by organization, their respective technology stacks, security baselines, and appetites for risk [4]. As organizations implement this architecture, and as with any technology endeavor, there inevitably comes a point of diminishing returns. The following questions automatically arise from this: How much microsegmentation is enough? How continuous does our "continuous monitoring" need to be?

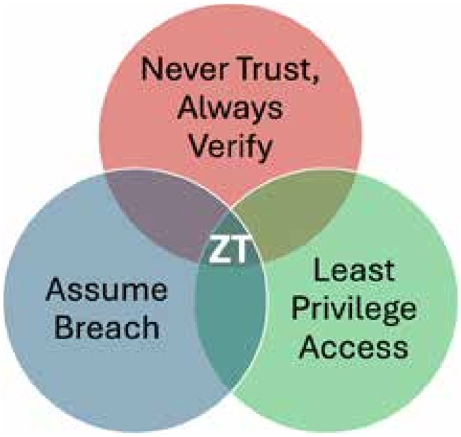


Figure 1. Key ZT Tenets *(Source: D. Bhasker and C. Zarcone).*

This article is a thought experiment where the authors explore a maximalist approach of the various ZTA philosophies to observe any trends that naturally determine the bookends of value generated by implementing ZTA.

BACKGROUND

From its origins in a doctoral dissertation by Stephen Paul Marsh and the "de-perimeterisation" work of the Jericho Forum to Forrester Research and Google's publication of its influential "BeyondCorp" series of white papers, ZTA has become mature by any architectural definition [5–8]. Many organizations are well on their way to ZTA, a process that typically spans multiple years and impacts almost every layer of the information technology stack. The process is often referred to as a "journey"—a recognition of the fact that for all but the smallest of organizations, migrating to ZTA will require sustained effort over a significant period.

> ***ZTA has spawned architectures and implementations that vary by organization, their respective technology stacks, security baselines, and appetites for risk.***

Several ZTA and ZTA-compatible frameworks have proliferated over time. The differences between these frameworks often reflect the philosophies of its developers, the priorities of organizations for which they are intended, and any procedural or technical constraints. Still, they tend to share a common set of characteristics, including the following:

- De-emphasis of networks as trust factors
  - the de-perimeterization of computer network boundaries
  - enhanced emphasis on small, workload-specific perimeters (often called "microsegmentation")
- Strong user and device authentication
- User and device policy compliance
- Continuous visibility and risk assessment of users and devices

All of this begs an interesting thought experiment—what happens if we extend the basic concepts of ZTA to its logical extremes? For example, how far do we go with de-perimeterization? Do we just eliminate perimeters altogether? How frequently do we authenticate our users—frequently to the point of continuously? How much compliance is enough? Do we allow a margin of lenience or mandate zero deviation from policy?

THE DISAPPEARANCE OF THE EDGE

The information security community has long used firewalls and other perimeter controls to segregate networks of different trust levels. But if one takes a maximalist view, ZTA suggests that all computer networks should be equally untrusted. From that perspective, it does not make sense to segregate networks of identical security posture; this is akin to building a fence in the middle of a cornfield—separating the corn from more corn.

If identity is the new perimeter, then why bother with the old? Do away with the edge altogether—fill in the castle moat, drop the drawbridge, and focus on fortifying the castle's keep with elite guards.

This approach is ideologically consistent. But at the same time, it does not make much practical sense since trust and control are two different things. ZTA teaches that networks should not be trusted but their controls can still be used to apply security policy, such as screening out undesirable traffic. An IP address might not be trusted for authentication purposes, but using a negative IP reputation score to inform real-time risk calculations is welcomed.

Despite the mantra that "the perimeter is dead," edge controls still add a level of value, even in the ZTA world. At a minimum, they provide a level of pest control, keeping the ScRiPt KidDiEz of the world at bay. They also serve as strategic choke points for monitoring and response. But their effectiveness is inversely proportional to their size, hence the movement toward microperimeters and workload-specific segmentation. End-user enclaves like home networks are also served well by perimeters, given their relatively small sizes and fixed geographic locations.

GOOD FENCES MAKE GOOD NEIGHBORS

In the ZT paradigm, the new perimeter is not at the edge of a network; it is deep into the interior, at the level of individual workloads. Creating small, manageable perimeters designed to protect a distinct application or set of applications under common administration is the new network security ideal. This is often called microsegmentation. Like watertight compartments of a seafaring ship's hull, the goal of microsegmentation is to isolate and limit the damage that occurs when any individual compartment is breached (Figure 2).

How segmented should a microsegmented network perimeter be? There are many possibilities, such as the following:

- Aggregate – Network policy could be applied to the entire address space assigned to a workload; perhaps a /22 of IPv4 address space or a /56 of IPv6.
- Subnet – The aggregate could be further subdivided into several subnets as necessary, with a tailored network policy applied to each subnet.

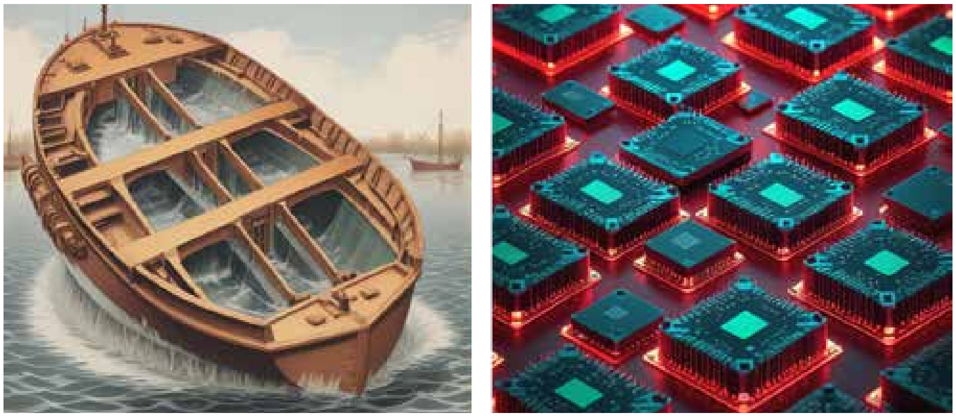


Figure 2. Ship Compartmentalization Compared With Microsegmentation Isolating Breach *(Source: Weinstein et al. [9])*.

> ***Creating small, manageable perimeters designed to protect a distinct application or set of applications under common administration is the new network security ideal.***

- Server – Network policy could be defined down to the level of individually addressed servers, regardless of subnet or aggregate membership.
- Operating System – Mainly by using host-based firewalls, individual servers could apply network policy to the individual services bound and listening on that server.
- Process – Individual processes on the server (native or containerized) could apply their own network security policies.

Any of these levels could serve as adequate microperimeters. However, they could also be applied in series, yielding a layered approach. So, how many layers are enough, and how many are too many?

"But wait a minute," someone may say, "I thought ZTA said that all networks are untrusted and that perimeters are passé, and here we are, building more perimeters. What gives?"

Once again, trust factors should not be confused with controls and control objectives. Even though the ship's hull might have watertight compartments to enhance buoyancy (microsegmentation), the ship cannot do away with carrying lifeboats, fire extinguishers, or communicating with vessel traffic controllers for navigation updates. ZTA can work cohesively with existing controls.

WHO ARE YOU? WHO ARE YOU AGAIN?

ZTA philosophy displaces networks as trust factors and, in turn, places stronger emphasis on user and device identity. Users and devices establish identity via authentication, typically multifactor authentication, for interactive user authentication and cryptographically strong mechanisms for devices.

ZTA calls for strong authentication to initiate a work session. Optionally, reauthentication can be required in situations where the risk profile of a user or device suddenly changes. But why use a reactive stance? Why not proactively reauthenticate a user or device every hour? Why not every 30 minutes, 10 minutes, single minute, or subsecond?

Clearly, users do not want to be impacted with authentication fatigue—as such, continuous interactive authentication is not an option. But the increasing use of second factors—user and/or device digital certificates installed on devices, hardware-based authenticators like FIDO 2 tokens, and passkey-enabled smartphones—raises the possibility of perpetual authentication. A service endpoint—whether an application or a VPN concentrator or an identity-aware access proxy—could interrogate such authenticators frequently, with no user intervention and negligible resource impact. Failed authentication (or the absence/removal of the authenticators) could have several explanations—it could be an innocuous change, such as loss of network connectivity or the user going to lunch. Or the change might not be innocuous; perhaps a FIDO 2 token was physically removed. Who removed it? Why did they remove it? These occurrences could be used in calculations of contextual risk, which, in turn, could trigger policy-driven responses as needed.

It is true that perpetually challenging a digital certificate or interrogating a hardware token only confirms that those authenticators are still present—the status of the user remains unknown. But this is better than nothing. If nothing else, perpetual authentication can help build a more accurate risk profile for that device; the instant an authentication fails, something has changed.

SPEED DIALING 911

The "Assume Breach" tenet of ZTA presumes that the adversary is already in the environment. This places an organization in a state of proactive response—a state of constant vigilance across its people, processes, and technologies. Containment, eradication, and recovery become business-as-usual processes, even if there is no detectable incident at any given time.

Assume Breach is instated regardless of the state of the effectiveness of preventative controls such as "Secure by Design," threat modeling, code reviews, and other security hygiene best practices. These might include other ZTA tenets like policy compliance, microsegmentation, and continuous verification.

From a maximalist vantage point, an adversary in an organization's environment can and will use all attack methods, tactics, techniques, and procedures at their disposal, with potentially all exploitation objectives (from reconnaissance, disruption of service, and data exfiltration to malware, ransomware, and Zero-Day exploits) to maximize damage. This, in turn, would imply that defensive security is constantly operating in a state of detecting, containing, and recovering from all possible breaches simultaneously. Exercising all security remediation, recovery technologies, and procedures would result in a self-directed assault on the environment. As the phrase implies, the dose of the medicine makes it a cure or poison.

Not unlike the villagers who got weary of rushing to the wolf boy's rescue, Security Operations Center (SOC) fatigue is real. It is reported on average that SOC deals with a 20% rate of false positives, with experts spending a third of their time on incidents that do not pose threats to the organization [10, 11]. SOC fatigue results in experts missing the subtle indicators of threat actors' presence in the environment.

> ***It is reported on average that SOC deals with a 20% rate of false positives, with experts spending a third of their time on incidents that do not pose threats to the organization.***

Anomalous behavior from the baseline is not necessarily malicious. A remote login at an unusual hour could simply be a conscientious employee checking in on work during a family vacation. How much security expertise does an organization want to expend on such security events?

The Assume Breach tenet is a mindset that shifts the organization from focusing only on prevention of breaches or remediation after an incident has been detected (adversaries are often detected months after infiltration) toward continuous detection and recovery. However, establishing intelligent, rapid detection, response, containment, and recovery protocols must be measured. Breach response must be tactical, effective, and aligned with the protect surface and risk appetite of the organization. To optimize effectiveness, the following approaches are recommended:

- Prioritize "Protect Surfaces" (which are the smallest possible attack surfaces) for:
  - inspecting and logging all traffic before acting,
  - continuously monitoring all configuration changes, resource accesses, and network traffic for suspicious activity, and
  - establishing full visibility of all activity across all layers from endpoints and the network to enable analytics that can detect suspicious activity.
- Establish security baselines and gather ample contextual data to detect anomalies that flag threat actors from all deviations from the baseline.
- Practice "Table Top" war games and red-teaming exercises.
- Ensure robust detection and response playbooks with proper implementation.

HOW MUCH (ZERO) TRUST IN A ZETTABYTE?

While trust evaluation algorithms (Figure 3) are rarely explicitly called out as a core tenet of ZTA principles, they are central to each of them, and

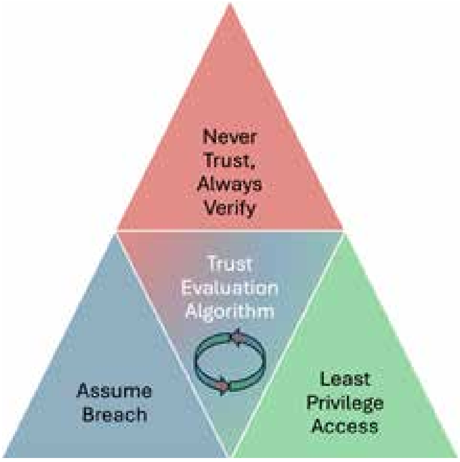


Figure 3. Trust Evaluation Algorithm *(Source: ACT-IAC [12]).*

they feed the processes of overall implementation (policy engines), automation, and orchestration.

However, this might also be the one Zero Trust parameter where more trust is better and maximum is best. Besides, how could one go wrong with maximum trust? The richer the telemetry, faster the analytics, and more fine-tuned the algorithm, the better the accuracy of risk calculations and context evaluations, thus arriving at reliable trust assertions that underpin ZTA policy decisions.

As the number of computers, servers, devices, and sensors continues to proliferate, they are generating large amounts of telemetry data (metrics, events, logs, and more). According to IDC's Global DataSphere forecasts, data generated from the core, edge, and endpoints are estimated to reach over 220,000 exabytes by 2026 [13, 14]. A myriad of issues needs to be dealt with in terms of data volumes, storage, normalization, cross correlation, streaming data, analytics, enrichment, privacy protections, and more. The richer the data, the greater the accuracy of the context and risk calculations for an entity. On the counterbalance, the greater the data volumes and associated processing, the more complexity in evaluating context. This gives rise to impedance with latency, lag, and errors in trust computations and, in turn, ZTA implementation.

These inefficiencies will drive trust algorithms toward finding a natural balance between the cost of computation, data storage, ingest, complexity, processing, and normalizing with the value of each datapoint in an algorithm used to assert trust. How much trust is enough will be determined by the risk appetite aligning with an organization's goals.

MATURITY ANALYSIS

Zero Trust is a journey—not a destination—and the journey is a complex one. If the Zero Trust approach falters, its cybersecurity benefits will significantly degrade [15]. Absolute Zero Trust does not exist and, in a sense, is not achievable for all the reasons discussed. Much of the "Extreme Zero Trust" discussion in this article illustrates this. The laws of diminishing returns are always in force, and Zero Trust is no exception.

> ***Data generated from the core, edge, and endpoints are estimated to reach over 220,000 exabytes by 2026.***

This is perhaps the main reason why frameworks like CISA's Zero Trust Maturity model culminate with the "optimal" state and not "absurd," "excessive," or "extreme."

ZTA may dismiss networks as trust factors, but this does not mean that network controls are entirely without value. In fact, many network controls are rolled into basic security hygiene to establish robust security baselines. Robust network tenancy and microsegmentation for applications and workloads are still desired but not unnecessary Protect Surfaces exposed to the internet. Factors like IP reputation scores are needed as part of intelligent risk analytics. Essentially, there is a shift in the way network controls are used in ZTA; however, they are still intrinsically valuable and remain a core part of an Optimal Network pillar.

Authentication is a security virtue. When applied too frequently, it becomes a hinderance to productivity and user acceptance, inevitably leading to workarounds. Depending on the architecture and use cases, it might be advisable to enable hardware-based

> ***Zero Trust is a journey—not a destination—and the journey is a complex one.***

authenticators to enable seamless polling to ensure continuous context awareness. Change of context or risk profile could trigger any number of policy responses like reauthentication. Essentially, factors optimizing productivity and user experience counterbalance the maximalist options for this tenet. Optimal identity is continuous and not confounding.

Assume Breach is a sensible position but should not succumb to paranoia and paralysis through analysis. If every log message and every signal is interpreted as a breach, the organization has bigger problems. False positives, security expertise expended on signals that do not threaten the organization, and SOC fatigue lead to lower quality of responses and depleted resources during true threats and malicious events. Optimization—knowing and prioritizing Protect Surfaces and efficiently analyzing signals and preparedness—is a must for Assume Breach tenet effectiveness. In the CISA model, "Visibility & Analytics" is depicted as a foundational overlay across the various pillars but is arguably its own pillar.

Lastly, the most precise trust algorithms require more data, decision points, and analysis—all potentially raising complexity in data analytics toward the point of diminishing returns. With data exponentially increasing in environments, keeping trust algorithms as streamlined as possible to arrive at an effective, valid trust decision is perhaps the best outcome for ZTA-centric policy decisions. Interestingly, data sits alone as both a pillar of protection ***and*** the asset protected by that pillar, creating a circular maturity dependency.

Through the course of the Zero Trust journey, proper risk management and alignment with technology strategy are essential. For example, organizations may choose to align with CISA's Zero Trust Maturity model. The risk profiles of those organizations will determine the ZTA measures implemented, and the technology strategy will drive the technical controls across the ZTA pillars (Identity, Devices, Networks, Applications, and Workloads and Data). As organizations journey toward the Optimal stage, they will find that they are at different levels of maturity across each pillar (Figure 4). Typically, it is challenging to advance across all pillars simultaneously. As a result, some pillars will be prioritized over others.

Moreover, since the ZTA tenets have interdependencies, developing one pillar requires constant vigilance in terms of cross-impacting to other pillars (or even other controls within the same pillar). For example, in the Data pillar, if encryption of data in transit (Initial), data at rest (Advanced), and data in use (Optimal) are all achieved, a data loss prevention

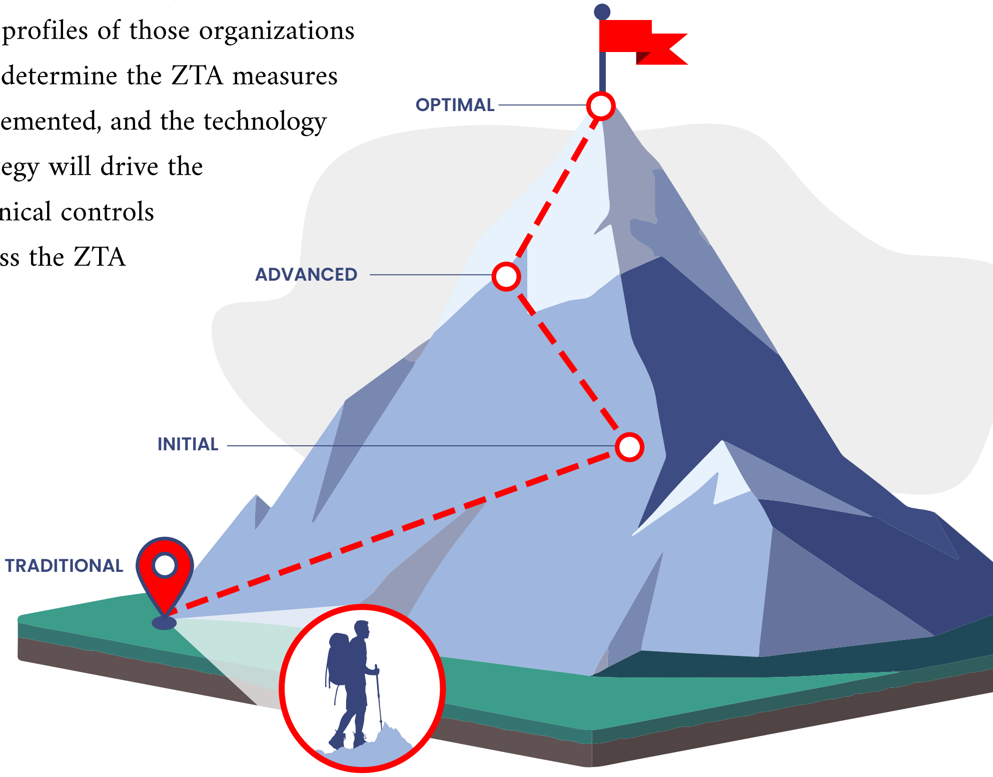


Figure 4. CISA's ZT Maturity Journey *(Source: CISA Cybersecurity Division [2]).*

control of the same pillar might lose visibility into the data and become a ghost control. Alternately, reaching levels of Optimal with encryption in the Data pillar might support Advanced maturity of the Network pillar by creating encrypted network flows.

Ultimately, ZTA maturity—in the CISA context or otherwise—involves a level of restraint. Optimal does not mean everything.

CONCLUSIONS

In review, when it comes to ZTA, more is certainly not better. Rather, what can be achieved is evaluating the ZTA tenets to find the right balance, considering an organization's unique goals, specific technology stack, Protect Surfaces, processes, assets, and culture with its multitude of specifics, that will determine the best-case scenario for each organization.

The Japanese concept of ikigai, the "reason for being"—that which gives an individual a sense of purpose—is found at the intersection of a confluence of factors (Figure 5). Similarly, the right measure of Zero Trust (which is not maximalist) is found at the intersection of a multitude of factors. Many of these factors are discussed deeply in publications such as NIST 800-207 and CISA's Zero Trust Maturity model. The Zero Trust Zen diagram (Figure 6) is rudimentary and for illustrative purposes only. It is not a comprehensive list of all possible factors that determine ZTA. ZTA is much more complex than the simplicity of life depicted in the ikigai diagram (Figure 5). As such, maximalist or extreme ZTA is not viable; in certain ways, it can be detrimental to the intent of Zero Trust principles. More is not necessarily better for the Zero Trust tenets explored. With that, one must ponder, what is your ikigai/what is your ZTA?

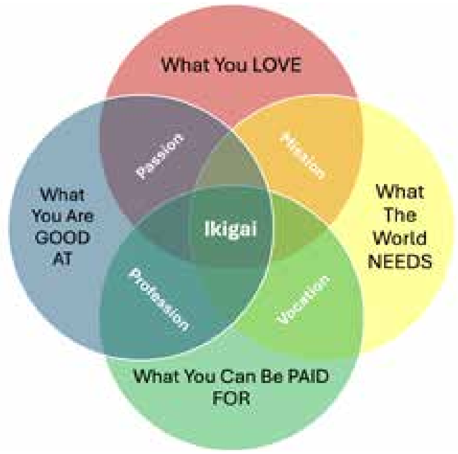


Figure 5. Ikigai: "Reason for Being" *(Source: Kaplan [16]).*

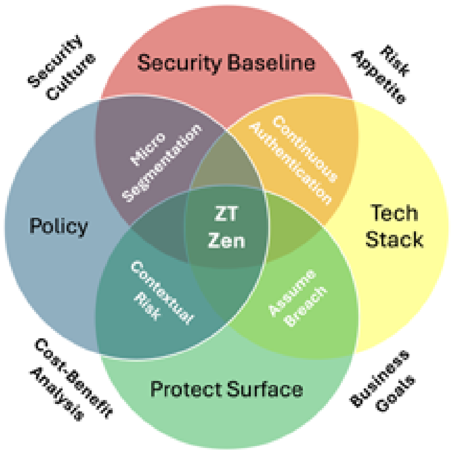


Figure 6. ZT Zen *(Source: Kaplan [16]).*

NOTE

Opinions expressed in this article are the authors' and not necessarily those of their employers. ■

REFERENCES

[1] National Institute of Standards and Technology (NIST) SP 800-207. "Zero Trust Architecture." https://csrc.nist.gov/pubs/sp/800/207/final, accessed 15 November 2023.

[2] Cybersecurity and Infrastructure Security Agency (CISA) Cybersecurity Division. "Zero Trust Maturity Model 2.0." https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf, accessed November 2023.

[3] Executive Order 14028. "Improving the Nation's Cybersecurity." 86 FR 26633, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/, accessed 10 December 2023.

[4] National Cyber Security Center. "Zero Trust Architecture Design Principles." https://www.ncsc.gov.uk/collection/zero-trust-architecture, accessed November 2023.

[5] The Open Group. "Secure Data: Enterprise Information Protection and Control." https://www.opengroup.org/sites/default/files/contentimages/Consortia/Jericho/documents/COA_EIPC_v1.1.pdf, accessed 23 November 2023.

[6] Ward, R., and B. Beyer. "BeyondCorp: A New Approach to Enterprise Security." https://storage.googleapis.com/gweb-research2023-media/pubtools/pdf/43231.pdf, vol. 39, no. 6, pp. 6–11, accessed 23 November 2023.

[7] Kindervag, J. "Build Security Into Your Network's DNA: The Zero Trust Network Architecture." *Forrester Research*, https://www.forrester.com/report/Build-Security-Into-Your-Networks-DNA-The-Zero-Trust-Network-Architecture/RES57047, accessed 23 November 2023.

[8] Marsh, S. P. "Formalising Trust as a Computational Concept." Ph.D. dissertation, University of Stirling, Stirling, Scotland, https://www.cs.stir.ac.uk/~kjt/techreps/pdf/TR133.pdf, accessed 23 November 2023.

[9] Weinstein, S., C. Zarcone, and S. Zevan. "Zero Trust Security Architecture for The Enterprise." SCTE Cable-Tec Expo 2022, https://www.nctatechnicalpapers.com/Paper/2022/FTF22_SEC04_Zarcone_3696, accessed 10 February 2024.

[10] Nadeau, J. "SOCs Spend 32% of the Day on Incidents That Pose No Threat." *Security Intelligence*, https://securityintelligence.com/articles/socs-spend-32-percent-day-incidents-pose-no-threat/, accessed 2 February 2024.

[11] "One-Fifth of Cybersecurity Alerts Are False Positives." *Security Magazine*, https://www.securitymagazine.com/articles/97260-one-fifth-of-cybersecurity-alerts-are-false-positives, accessed 5 February 2024.

[12] ACT-IAC. "Zero Trust Cybersecurity Current Trends." https://www.actiac.org/system/files/ACT-IAC%20Zero%20Trust%20Project%20Report%2004182019.pdf, accessed 2 February 2024.

[13] Burgener, E., and J. Rydning. "High Data Growth and Modern Applications Drive New Storage Requirements in Digitally Transformed Enterprises." *IDC*, https://www.delltechnologies.com/asset/en-my/products/storage/industry-market/h19267-wp-idc-storage-reqs-digital-enterprise.pdf, accessed February 2024.

[14] Reinsel, D., J. Gantz, and J. Rydning. "The Digitization of the World From Edge to Core." *IDC*, https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf, accessed 4 February 2024.

[15] National Security Agency. "Embracing a Zero Trust Security Model." https://media.defense.gov/2021/Feb/25/2002588479/-1/-1/0/CSI_EMBRACING_ZT_SECURITY_MODEL_UOO115131-21.PDF, accessed 19 January 2024.

[16] Kaplan, O. "Goals-vs.-Systems-and-Finding-Your-Ikigai." https://slashproject.co/posts/2020/goals-vs-systems-and-finding-your-ikigai/, accessed 20 November 2023.

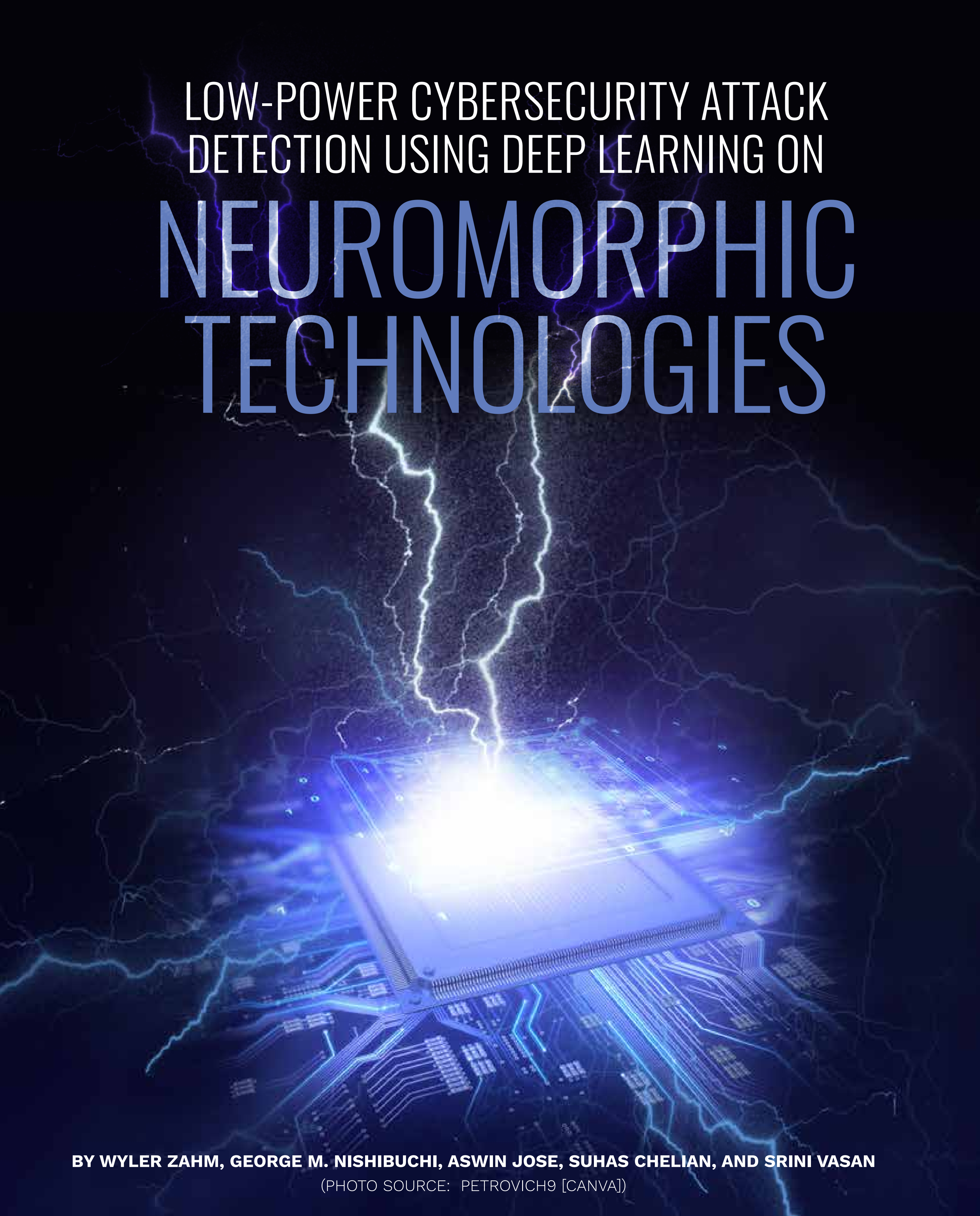
BIOGRAPHIES

DAKSHA BHASKER is a principal cybersecurity architect at Microsoft, where she works on M365 cloud security. Throughout her career, she has engaged with the broader security industry to better the state of security architectures of current and emerging technologies. She has over 20 years of experience in the telecommunications service provider industry, specializing in systems security development of complex solutions architectures. Ms. Bhasker holds an M.S. in computer systems engineering from Irkutsk State Technical University and an MBA in ecommerce from the University of New Brunswick, Canada.

CHRISTOPHER ZARCONE is a distinguished engineer at Comcast and a computer scientist specializing in Zero Trust. He has presented at information security conferences and seminars, contributed to the development of industry standards, and has been awarded 10 U.S. patents. He was recently named Engineer of the Year by IEEE's Philadelphia Section. Mr. Zarcone holds a bachelor's degree in computer science from Drexel University and a master's degree in computer science from Rensselaer Polytechnic Institute.

Photo Source: Billion Photos (Canva)

LOW-POWER CYBERSECURITY ATTACK DETECTION USING DEEP LEARNING ON NEUROMORPHIC TECHNOLOGIES



BY WYLER ZAHM, GEORGE M. NISHIBUCHI, ASWIN JOSE, SUHAS CHELIAN, AND SRINI VASAN

(PHOTO SOURCE: PETROVICH9 [CANVA])

SUMMARY

Neuromorphic computing systems are desirable for several applications because they achieve similar accuracy to graphic processing unit (GPU)-based systems while consuming a fraction of the size, weight, power, and cost (SWaP-C). Because of this, the feasibility of developing a real-time cybersecurity system for high-performance computing (HPC) environments using full precision/GPU and reduced precision/neuromorphic technologies was previously investigated [1]. This work was the first to compare the performance of full precision and neuromorphic computing on the same data and neural network and Intel and BrainChip neuromorphic offerings. Results were promising, with up to 93.7% accuracy in multiclass classification—eight attack types and one benign class.

Since then, a BrainChip Akida 1000 chip was acquired, and Intel released the Loihi 2 chip and developed the Lava framework for establishing neuromorphic deep-learning applications. These developments and more detailed analyses are reflected in this article, with up to 98.4% accuracy achieved in classifying nine classes. Compared to the state of the art, neuromorphic technologies have much smaller SWaP-C profiles. In addition, how these systems can be applied to deployable platforms like manned aircraft or unmanned aerial vehicles (UAVs) is discussed, and additional use cases of neuromorphic computing in computer vision are reviewed.

INTRODUCTION

In the work of Zahm et al. [1], the viability of a real-time high-performance computing (HPC)-scale cybersecurity system was evaluated using full precision/GPUs and reduced precision/neuromorphic technologies in a proof of concept called Cyber-Neuro RT (e.g., Figure 1). The government operates several HPC systems for the Defense Advanced Research Projects Agency (DARPA), Oak Ridge National Laboratory (ORNL), and National Aeronautics and Space Administration's (NASA's) Advanced Supercomputing Division. These systems operate at much larger scales than traditional information technology (IT) domains and thus require novel tools to address their unique requirements. Neuromorphic systems were investigated because they can achieve similar accuracy as GPU systems, with a fraction of the SWaP-C budgets.

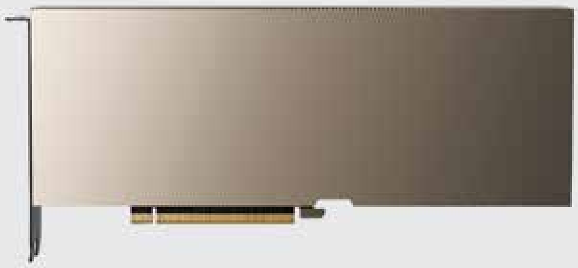
Zahm et al. [1] were one of the first to run the same data and neural network through full precision and Intel and BrainChip neuromorphic offerings. Intel results were generated with the Loihi 1 chip and the SNN-TB Toolbox. BrainChip results were generated via software simulation and the CNN2SNN Toolbox. Results were

Figure 1: Nanoscience High-Performance Computing System *(Source: Photo Courtesy of Argonne National Laboratory).*

promising—up to 93.7% correct on a dataset with 450,000 entries and nine classes, including eight attack types; however, further research was necessary. Since then, a BrainChip Akida 1000 chip was received, and Intel released the Loihi 2 chip and developed the Lava framework for programming it. These developments and more detailed analyses were reviewed, and improved accuracy was achieved on full precision/GPU and reduced precision/neuromorphic technologies, with up to 98.4% correct. This is comparable to the state of the art presented in Gad et al. [2] and Sarhan et al. [3]. However, neuromorphic technologies have the advantage of a much smaller SWaP-C envelope, as detailed in Table 1. Because neuromorphic processors achieve similar results to GPUs and with dramatic SWaP-C savings, they are well suited for embedded or edge

> ***Because neuromorphic processors achieve similar results to GPUs and with dramatic SWaP-C savings, they are well suited for embedded or edge devices.***

Table 1. Comparison of GPU and Neuromorphic Compute Platforms *(Source for Left Image, Wikimedia [https://commons.wikimedia.org/wiki/File:Nvidia_Tesla_A100.png]; Right, Intel)*

CPU/GPU PLATFORMS	SWaP-C	NEUROMORPHIC PLATFORMS
NVIDIA A100 PCIE 4.0 DUAL SLOT		USB KEY FORM FACTOR FOR INTEL LOIHI 1
		
26.7 L x 11.2 H x 3.5 W	Size (cm)	5.1 L x 1.3 H x 0.6 W
1674	Weight (g)	~180
250	Power (W)	1
~$16,000	Cost (USD)	$50 (est.)

Note: Because Intel has not released packaging information on Loihi 2, only Loihi 1 is listed in this table.

devices. SWaP-C savings for neuromorphic computing could be even better when using field-programmable gate arrays (FPGAs) or a licensable Intellectual Property in Register-Transfer Level (IP cores in RTL) format.

METHODS

Datasets

A subset of the University of New South Wales (UNSW) TON-Internet of Things (TON-IoT) dataset was used [4]. Commonly used by cybersecurity researchers, this dataset has nine attack classes listed in Table 2 but excluding the man-in-the-middle (MITM) class due to the small sample count. Several other datasets were investigated, but only the TON-IoT was included in this work. The data-cleaning and feature-encoding process was an extension of Zahm et al. [1]. This process now infers data types more reliably, including supporting list type features such as the "all_headers" field, which contains a list of all headers of an HTTP communication using Zeek's http.log. After examining the data by packet structure and excluding timestamp and unique identifier, deduplication and deconfliction were found to be useful. In deduplication, identical data samples were removed. In deconfliction, identical data samples with different labels were removed. This process mitigates model overfitting. After deduplication and deconfliction, ~300,000 samples remained.

Table 2. TON-IoT Dataset Summary

CLASS	TRAINING	TESTING
Normal	66,916 (~43%)	16,729 (~43%)
DDOS	11,120 (~7%)	2,780 (~7%)
DOS	11,120 (~7%)	2,780 (~7%)
Password	11,120 (~7%)	2,780 (~7%)
Scanning	11,120 (~7%)	2,780 (~7%)
XSS	11,120 (~7%)	2,780 (~7%)
Injection	11,120 (~7%)	2,780 (~7%)
Backdoor	11,120 (~7%)	2,780 (~7%)
Ransomware	11,120 (~7%)	2,780 (~7%)
Total	155,876 (100%)	38,969 (100%)

The MITM class was merged into injection attacks data due to the similarity of the two attacks and low MITM data count (0.03% of the original data). The number of remaining samples was then balanced for attack classes, and the remainder was left for normal traffic. This balance yielded a total training dataset size of 194,845 rows. Although smaller than the final Zahm et al. [1] dataset, it has proportionately less normal traffic (42.9% vs. 66%). Deduplication prevented skewing accuracy figures up by repeatedly processing the same data point and ensuring the test split did not contain samples found in the train set (e.g., if the data point is accurately classified, accuracy becomes inflated). Table 2 shows the distribution of class types where an 80%/20% train/test split was used. Chance performance would be to always declare the majority class of normal and yield 42.9% accuracy.

Full-Precision Neural Network and Hyperparameter Tuning

A fully connected neural network was used for classification. In prior work, encoder units (but not decoder units) of a separately trained autoencoder network combined with a multilayer, fully connected network were used. This resulted in a 10-layer, fully connected neural network (one input layer, four autoencoder layers, and five network layers). The autoencoder was removed due to its highly variable post-training performance across different datasets. This also improved model compatibility with the BrainChip hardware. After spiking neural network (SNN) conversion, Zahm et al.'s model [1] required more than 80 neural processing units. This was over the limit for Akida's neuron fabric, so only software simulation could be performed. By removing the autoencoder, the model parameter count was reduced, enabling the model to run on the BrainChip's hardware and allowing timing and power data to be measured.

The "Weights and Biases" grid search [5] was also used to identify optimal hyperparameter values like learning rate, batch size, feature count, and hidden layer sizes. Feature count refers to the number of features chosen from the dataset, ranked by feature importance from a random forest model. Too few features may lead to underfitting, while too many may lead to overfitting. Similarly, as hidden layer widths increase, a risk of overfitting arises, and using too few neurons may underfit the data.

BrainChip Neuromorphic Platform

BrainChip offers the Akida 1000 chip for edge artificial intelligence/machine-learning (AI/ML) applications. These chips perform the inference steps of SNNs. BrainChip provides the CNN2SNN Python toolkit [6] to assist in converting full-precision artificial neural network (ANN) models to Akida-compatible SNN models. Zahm et al. [1] utilized the CNN2SNN toolkit to perform a simple ANN-SNN conversion process, with room for improvement. In this work, the following were explored: (1) the effects of an improved input data scaling process, (2) improved ANN to SNN conversion process, (3) running the converted model on Akida hardware, and (4) measuring model performance, power, and timing on the device.

> ***BrainChip offers the Akida 1000 chip for edge artificial intelligence/machine-learning (AI/ML) applications.***

Input Data Scaling

Data scaling is crucial to optimize SNN performance and needs to be tuned by the dataset and model. Data scaling for SNNs modifies traditionally zero-centered and unit standard deviation scaled data such that its maximum, minimum, and spread optimally cover the target input range. For the Akida system on a chip, this input range is limited to four unsigned bits, or [0, 15], by the hardware.

SNNs have 4- to 8-bit neuron firing rates compared to floating point activations in traditional ANNs. Data must be scaled appropriately to the range of the limited neuron firing rates to maximize the information passed through the network. In Zahm et al. [1], the data was rescaled from 0 to 1. To make it compatible with 4-bit precision, data was then multiplied by 10 and rescaled to a 4-bit unsigned integer.

For this work, outliers outside 95% of the upper range of data and 5% of the lower range were removed and min-max rescaling applied. Versus Zahm et al. [1], this produced a more uniform distribution of the data across the 4-bit range. For example, feature id.resp_p_conn ranged in [0, 8] with the old scaling method but [0, 15] with the new method. The scaled data was then binned to the nearest whole integer in the range. The binned data was divided by 15 to the [0, 1] range, and quantization aware retraining was applied. The final on-chip model inference function was passed scaling factors (x - ab), along with the binned, 4-bit unsigned data.

The effect of data scaling on quantized models and converted Akida SNNs was measured. (Quantization refers to reducing the bit width of an ANN model's weights and biases from higher to lower precision.) Conversion takes a quantized model and removes training-only features like dropout.

ANN to SNN Conversion

SNNs have lower precision representations in their weights and activation functions compared to ANNs. This allows for simpler hardware due to the removal of floating-point computations—this is one reason why neuromorphic networks use less power with smaller models. However, simply changing the precision of model weights and scaling appropriately does not produce performant models.

Zahm et al. [1] developed the following quantization process: quantize weights to a high, unsigned precision; execute quantization-aware retraining with the original training data; and then convert to a lower precision. Retraining continued until learning plateaued. In this work, more quantization steps were used. The quantization schedule is shown in Table 3. These hyperparameters were determined empirically through grid search. Future exploration could introduce greater granularity into this analysis.

In addition, a monotonically decreasing learning rate schedule was used at each quantization step. Learning rates decreased by a factor of 10 over four steps, thus allowing additional training iterations for fine tuning. The training schedule was set to minimize the number of experiments over the large search space. Future work could use Bayesian methods to determine learning rate schedules for different network architectures or datasets.

Table 3. Quantization Schedule

STEP	WEIGHT PRECISION	ACTIVATION PRECISION
0	32-bit floating point	32-bit floating point
1	6-bit unsigned	4-bit unsigned
2	4-bit unsigned	4-bit unsigned

On-chip Execution

Zahm et al. [1] leveraged the BrainChip software simulator for SNN execution. This work uses the BrainChip Akida device to analyze SNN model performance.

Intel Neuromorphic Platform

This article focuses on Intel's latest Loihi 2 chip with their new deep-learning framework, Lava Deep Learning (Lava DL) [7]. Zahm et al. [1] utilized the Loihi 1 architecture combined with the SNN-Toolbox, a software package intended for direct ANN-SNN conversion. The Lava DL software package contains two main modules for training networks compatible with Loihi hardware—Spike Layer Error Reassignment in Time (SLAYER) and Bootstrap. SLAYER is intended for native training of deep event-based networks, and Bootstrap is intended for training rate-coded SNNs.

The Bootstrap module of the Lava-DL accelerates the training of SNNs and closes the performance gap compared to an equivalent ANN. SNNs have extended training times compared to ANNs. This method leverages the similarity between the behavior of the leaky integrate and fire (LIF) neuron and the rectified linear unit activation function to produce a piecewise mapping of the former to the latter. This method of ANN-SNN syncing in training is particularly beneficial because it accelerates the training of a rate-coded SNN, reduces the inference latency of the trained SNN, and closes the gap between ANN and SNN accuracy. The network was trained on either a CPU or GPU, and then inference was performed on the Loihi 2 hardware. Testing on an identically structured network to the Akida hardware tests was performed to compare performance between the two product offerings.

The current-based LIF (CUBA) neuron model, combined with an Adam optimizer, and categorical cross entropy loss were used to train the model. The neuron threshold parameter of the LIF neuron affected performance the most of any of the neuron parameters, so values from 0.25 to 1.5 in steps of 0.25 were tested. If the threshold was too low, performance would suffer due to saturation of neuron activation in the

> ***The neuron threshold parameter of the LIF neuron affected performance the most of any of the neuron parameters.***

subsequent layers. If the threshold was too high, few neurons would activate at all and the performance of the network would suffer. A value of 0.75 provided the best performance. Altering the default values of the other parameters induced erratic neuron behavior and unstable training.

Hyperparameter Tuning

After identifying a sufficient neuron model for the Lava DL model, further hyperparameter optimization was performed for the batch sizes of 256, 512, and 1,024 transactions and learning rates of 1E-3, 1E-4, and 1E-5. Training was performed for 200 epochs for each model to allow convergence for the lowest learning rates.

RESULTS

Full-Precision Neural Network and Hyperparameter Tuning

While all values had an impact on model performance, hidden layer count and batch size were of greatest value for maximizing model performance. Plotting the average of a performance metric for all sweeps grouped by the parameter of interest also shows these relationships. Increased batch size was correlated with decreased model test accuracy (see Figure 2), and increased parameter count via hidden unit design correlated positively with accuracy (not shown). Larger batch sizes sometimes led to more unstable training and decreased accuracy.

By hyperparameter tuning across more than 50 parameter configurations, a model accuracy of 98.42% was achieved on the TON-IoT subset split data with an 80/20 split. This parameter sweep was performed on two NVIDIA A100 nodes and took three hours. Previous hyperparameter sweeps took more time due to larger neural network sizes.

BrainChip Neuromorphic Platform

The following results are presented in this section: an improved data scaling process, an improved ANN to SNN conversion process, and running the converted model on a chip rather than just via software simulator. This gained valuable insights into real hardware inference speeds and power costs.

Data Scaling

The new data scaling method outperformed the old Zahm et al. [1] method on both quantized and converted SNN models, as shown in Table 4. Quantized model performance improved ~3.1% and converted SNN model performance improved ~5.3%. Log-scaling was also tried but did not perform as well as the new method. The reduced accuracy drop was noted when converting the quantized model to an

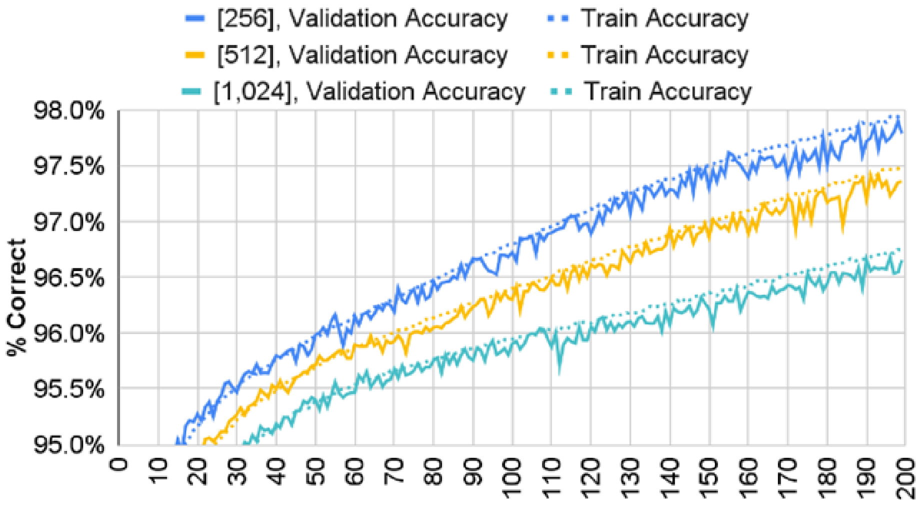


Figure 2. Sweep Mean Train (Dotted Line) and Validation Accuracy (Solid Line) for Different Batch Sizes *(Source: Zahm et al.).*

Table 4. BrainChip, SNN Data Scaling Technique vs. Performance

SCALING METHOD	ANN % ACCURACY	QUANTIZED % ACCURACY	CONVERTED % ACCURACY
Zahm et al. [1]	94.66	87.92	83.98
Ours	94.66	91.00	89.30

SNN with the data scaled via the new processes (3.9% vs. 1.7%). Passing scaling factors were also introduced to the SNN model in hardware to further reduce this accuracy loss when converting a quantized model. However, this was not used to produce Table 4, as Zahm et al. [1] used this technique. Data scaling experiments were done for SNN training but not ANN training, hence identical ANN performance. Note that identical initial ANN model and identical quantization retraining schedules were used and not the optimal ANN design and optimal quantization retraining schedules.

ANN to SNN Conversion

In Table 5, quantization yields dramatically smaller models to fit on low SWaP-C neuromorphic hardware. Accuracy increased for the ANN and SNN models, and ANN performance increased 4.7%. At 98.4% accuracy, this was similar to the state of the art presented in Gad et al. [2] and Sarhan et al. [3]. Improvements to the quantization schedule reduced the accuracy drop from 11.2% to 7.2% between full and reduced precision models.

On-chip Execution

Results for ANN vs. BrainChip SNN size, power, and speed are summarized in Table 6. Power consumption was ~1 W. GPU power was estimated at 30 W, using 10% of an NVIDIA A100's maximum power consumption. Speed was slower for neuromorphic chips. GPU models could operate with much higher throughput due to batch processing, which might not be available for streaming cybersecurity data.

Intel Neuromorphic Platform

Batch size was negatively correlated with accuracy, while learning rate was positively correlated with accuracy. Larger batch sizes took longer to converge but were less susceptible to random fluctuations in the dataset. The Bootstrap framework appeared to perform better with larger learning rates, whereas ANNs typically preferred smaller learning rates.

Table 5. BrainChip, Accuracy Benchmarks

METHOD	BIT WIDTH	ANN % ACCURACY	SNN % ACCURACY
Zahm et al. [1]	8	93.7	82.5
New method	4	98.4	91.2

Table 6. ANN vs. BrainChip SNN Size, Power, and Speed

MODEL	DEVICE	SIZE (KB)	POWER USAGE	SPEED
ANN	NVIDIA A100	157	30 W*	29,412 Hz
SNN	Akida 1000	15	929 mW	5,104 Hz†

* Device characteristics provided instead at an estimated 10% of max power usage for model inference.

† Average over 30 trials due to variability in single executions.

A final accuracy of 90.2% was achieved with the Lava DL Bootstrap framework, with an identical architecture to the Akida network, as shown in Table 7. This was a reduction in accuracy of 3.5% compared to the prior work of Zahm et al. [1]. However, the old SNN-Toolbox performed direct ANN-SNN conversion, while Lava DL required implementation and training of a native SNN.

A 72.4% reduction in model size was observed between the full-precision ANN and the Lava DL model detailed in Table 8. With over 24 MB of memory available on Loihi 2 chips, this model is expected to comfortably fit on the hardware.

While the Lava DL network could predict normal traffic 99% of the time, it struggled to accurately predict the precise class of non-normal traffic. The highest classification accuracy in non-normal traffic was 52% for DOS attacks.

DISCUSSION

A final accuracy of 90.2% was achieved with the Lava DL Bootstrap framework, with an identical architecture to the Akida network.

The following was presented from this work: an improved dataset with less normal traffic and improved ANN performance via better data preprocessing and hyperparameter tuning. For BrainChip, accuracy improved, model size decreased, and the model on the Akida chip was assessed for timing and power. Improvements were attributed to better data scaling and rigorous model quantization and retraining. For Intel, the performance of the new Lava DL framework was benchmarked, with a slight dip in performance compared to the prior SNN-Toolbox. However, accuracy was similar to BrainChip. Although the percentage of correct results (~98%) was like the state of the art presented in Gad et al. [2] and Sarhan et al. [3], low neuromorphic processors could be used with dramatic SWaP-C savings (see Table 1). In related work, a semi-supervised approach to cybersecurity on Intel's Loihi 2 was investigated [8]. Testing these models on Intel hardware and larger and more diverse datasets is a goal for future work.

CONCLUSIONS

Because of their low SWaP-C envelope, neuromorphic technologies are well suited for deployable platforms such as manned aircraft or UAVs. Table 1 illustrates the SWaP-C advantages of neuromorphic processors compared to GPUs. Neuromorphic technologies could be used for cybersecurity of embedded networks or other functions like perception or control. Network traffic across CAN buses, for example, could be passed through neuromorphic processors. These processors would then detect abnormal traffic, which could then be blocked, preventing further harm.

Neuromorphic computing was also pursued for computer vision projects. Park et al. [9] used an ANN to SNN conversion to classify contraband materials across a variety of conditions, such as different temperatures, purities, and backgrounds. Neuromorphic technologies for image processing and

Table 7. Intel Accuracy

METHOD	% ACCURACY
Zahm et al. [1]	93.7
New method	90.2

Table 8. ANN vs. Intel Size and Speed

METHOD	ANN SIZE	SNN SIZE	SPEED
Zahm et al. [1]	19.7 MB	6.5 MB	500 Hz (Loihi 1 hardware)
New method	157 KB	44.4 KB	9090 Hz (Loihi 2 simulation)

> ***Because of their low SWaP-C envelope, neuromorphic technologies are well suited for deployable platforms such as manned aircraft or UAVs.***

automatic target recognition were also explored [10]. For image processing, hierarchical attention-oriented, region-based processing (HARP) [11] was used. HARP removes uninteresting image regions to speed up image transfer and subsequent processing. For automatic target recognition, U-net was used to detect tiny targets in infrared images in cluttered, varied scenes. U-net was run on Intel's Loihi chip [12].

Research and development in cybersecurity and neuromorphic computing continues, with great potential in both.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research program under Award Number DE-SC0021562. ■

REFERENCES

[1] Zahm, W., T. Stern, M. Bal, A. Sengupta, A. Jose, S. Chelian, and S. Vasan. "Cyber-Neuro RT: Real-time Neuromorphic Cybersecurity." *Procedia Computer Science*, vol. 213, pp. 536–545, 2022.

[2] Gad, A., A. Nashat, and T. Barkat. "Intrusion Detection System Using Machine Learning for Vehicular Ad Hoc Networks Based on ToN-IoT Dataset." *IEEE Access*, vol. 9, pp. 142206–142217, 2021.

[3] Sarhan, M., S. Layeghy, N. Moustafa, M. Gallagher, and M. Portmann. "Feature Extraction for Machine Learning-Based Intrusion Detection in IoT Networks." *Digital Communications and Networks*, vol. 10, pp. 205–216, 2022.

[4] Moustafa, N. "New Generations of Internet of Things Datasets for Cybersecurity Applications Based on Machine Learning: TON_IoT Datasets." In Proceedings of the eResearch Australasia Conference, Brisbane, Australia, pp. 21–25, 2019.

[5] Weights and Biases. "Weights and Biases." http://www.wandb.com, accessed on 14 April 2023.

[6] BrainChip. "CNN2SNN Toolkit." https://doc.brainchipinc.com/user_guide/cnn2snn.html, accessed on 14 April 2023.

[7] Intel. "Lava Software Framework." https://lava-nc.org/, accessed 14 April 2023.

[8] Bal, M., G. Nishibuchi, S. Chelian, S. Vasan, and A. Sengupta. "Bio-Plausible Hierarchical Semi-Supervised Learning for Intrusion Detection." In Proceedings of the International Conference on Neuromorphic Systems (ICONS), Santa Fe, NM, 2023.

[9] Park, K. C., J. Forest, S. Chakraborty, J. T. Daly, S. Chelian, and S. Vasan. "Robust Classification of Contraband Substances Using Longwave Hyperspectral Imaging and Full Precision and Neuromorphic Convolutional Neural Networks." *Procedia Computer Science*, vol. 213, pp. 486–495, 2022.

[10] SBIR.gov. "Bio-inspired Sensors." https://www.sbir.gov/node/2163189, accessed 14 April 2023.

[11] Bhowmik, P., M. Pantho, and C. Bobda. "HARP: Hierarchical Attention Oriented Region-Based Processing for High-Performance Computation in Vision Sensor." *Sensors*, vol. 21, no. 5, p. 1757, 2021.

[12] Patel, K., E. Hunsberger, S. Batir, and C. Eliasmith. "A Spiking Neural Network for Image Segmentation." arXiv preprint arXiv:2106.08921, 2021.

BIOGRAPHIES

WYLER ZAHM is a researcher and senior ML engineer. He has worked with advanced algorithms, front- and back-end development, a variety of AI/ML architectures and frameworks such as full precision/GPU and reduced precision/neuromorphic technologies, and applications like automated vulnerability detection and repair for computer source code and cybersecurity. Mr. Zahn has dual bachelor's degrees in computer engineering and data science from the University of Michigan.

GEORGE NISHIBUCHI is a researcher in materials science and DL. He has a background in computational materials science, with experience running over 50,000 Density Functional Theory simulations at Purdue University's Network for Computational Nanotechnology, including phonon studies of infrared transparent ceramics, high-throughput studies of semiconductors, and mechanistic studies in solid-state electrolytes. He has also contributed to research in neuromorphic learning algorithms for network intrusion detection systems. Mr. Nishibuchi has an M.S. in materials engineering from Purdue University.

ASWIN JOSE is a lead system engineer with more than 12 years of extensive experience in system design, software architecture, and leadership. He possesses a broad and deep expertise in various domains such as generative AI/ML verification and validation (V&V), computer vision, big data analytics, the banking sector, logistics, semiconductor technology, healthcare systems, and cutting-edge technological frameworks like full-stack architectures, lambda architecture, and the MEAN stack. Mr. Jose holds an M.E. in computer science from Anna University.

SUHAS CHELIAN is a researcher and ML engineer. He has captured and executed more than $12 million worth of projects with several organizations like Fujitsu Labs of America, Toyota (Partner Robotics Group), HRL Labs (Hughes Research Lab), DARPA, the Intelligence Advanced Research Projects Agency, and NASA. He has 31 publications and 32 patents demonstrating his expertise in ML, computer vision, and neuroscience. Dr. Chelian holds dual bachelor's degrees in computer science and cognitive science from the University of California, San Diego and a Ph.D. in computational neuroscience from Boston University.

SRINI VASAN is the president and CEO of Quantum Ventura Inc. and CTO of QuantumX, the research and development arm of Quantum Ventura Inc. He specializes in AI/ML, AI V&V, ML quality assurance and rigorous testing, ML performance measurement, and system software engineering and system internals. Mr. Vasan studied management at the MIT Sloan School of Management.

NAVIGATING CHALLENGES AND OPPORTUNITIES IN THE CYBER DOMAIN WITH

SIM2REAL TECHNIQUES

BY EMILY A. NACK AND NATHANIEL D. BASTIAN

(PHOTO SOURCE: ZABELIN [123RF.COM])

INTRODUCTION

In the digital age, the cyber domain has become an intricate network of systems and interactions that underpin modern society. Sim2Real techniques, originally developed with notable success in domains such as robotics and autonomous driving, have gained recognition for their remarkable ability to bridge the gap between simulated environments and real-world applications. While their primary applications have thrived in these domains, their potential implications and applications within the broader cyber domain remain relatively unexplored. This article examines the emerging intersection of Sim2Real techniques and the cyber realm, exploring their challenges, potential applications, and significance in enhancing our understanding of this complex landscape.

SIM2REAL: CONCEPTS AND METHODOLOGIES

Sim2Real, an abbreviation for “Simulation to Reality,” is a transformative approach that addresses the challenge of transferring knowledge acquired in simulated environments to real-world applications [1]. It plays a vital role in various domains by leveraging simulated environments to

TABLE OF CONTENTS

> ***Sim2Real techniques have gained recognition for their remarkable ability to bridge the gap between simulated environments and real-world applications.***

train and prepare for real-world scenarios. While Sim2Real is commonly associated with machine learning (ML), its applications extend beyond this field, offering opportunities for enhanced learning, testing, and preparation in diverse scenarios. This section explores Sim2Real's foundational concepts and methodologies, which collectively enable the seamless transfer of knowledge from simulation to practical applications, a principle that has wide-ranging implications, including potential applications within the cyber domain.

The Foundation of Sim2Real

At the heart of Sim2Real lies the concept of training ML models in simulated environments, where data is abundant, diverse, and controllable. This approach stands in contrast to traditional methods, which often require training models directly in real-world settings, where data collection can be expensive, limited, or impractical. By leveraging the advantages of simulation, Sim2Real techniques enable the rapid development, refinement, and evaluation of ML models, offering a more cost-effective and flexible solution.

The adoption of Sim2Real techniques is rooted in addressing the limitations and challenges associated with real-world model training. These challenges can be broadly grouped into the following key areas, as defined in references 2–7 and highlighted in Table 1.

Data Abundance and Diversity

Real-world data collection often falls short in providing diverse and ample datasets. This limitation can hinder the training of ML models, particularly in tasks requiring robust generalization. In Sim2Real, simulations serve as a solution. These environments offer a vast and diverse source of data, creating extensive training datasets [2, 3]. Within these digital realms, researchers have the capacity to introduce an expansive array of scenarios, from varied terrains to different weather conditions and physical properties [5]. The diversity inherent in simulations empowers models to generalize effectively, adapting seamlessly to a wide spectrum of real-world scenarios [7].

Cost-Effectiveness and Flexibility

The cost of data collection in real-world settings can be a significant

Table 1. Comparative Analysis of Traditional Model Training vs. Sim2Real Techniques

FACTORS	TRADITIONAL METHODS	SIM2REAL TECHNIQUES
Data Abundance and Diversity	Limited and costly real-world data collection	Abundant and diverse simulated datasets. Researchers can introduce various scenarios, including different terrains, weather conditions, and physical properties, creating extensive training datasets.
Cost-Effectiveness	Expensive data collection in real-world settings, requiring specialized equipment, personnel, and logistics	Cost-effective due to reduced expenses linked to data collection. Sim2Real accelerates model development and provides unparalleled flexibility, facilitating rapid prototyping and iterative model development.
Flexibility	Limited experimentation in real-world scenarios	Highly flexible, with the ability to experiment effortlessly with various scenarios. This adaptability contributes to more efficient and agile model design.
Safety and Risk Mitigation	Real-world testing carries risks to humans and equipment	Enhanced safety in simulated environments. Researchers can engage in complex experimentation without risk to human operators or expensive equipment.

barrier, particularly in domains such as autonomous vehicles. Traditional methods may not only be expensive but also less flexible when it comes to experimenting with different scenarios [4, 7]. Sim2Real, in this context, emerges as a transformative solution. This approach substantially reduces expenses linked to data collection, making it a crucial factor for industries where cost efficiency is imperative. Sim2Real's cost-effective nature accelerates model development and, more importantly, provides unparalleled flexibility. Researchers can effortlessly experiment with various scenarios, facilitating rapid prototyping and iterative model development [5]. This adaptability contributes to more efficient and agile model design.

Safety and Risk Mitigation

Safety is a paramount concern in high-stakes domains like healthcare, aerospace, and disaster response. Real-world testing in these areas can carry significant risks to human operators and valuable equipment [6]. Simulated environments emerge as the safer alternative [2–4]. Within these controlled digital realms, the safety of both human operators and valuable assets is a top priority. Sim2Real effectively mitigates the risks associated with real-world testing. Researchers can engage in complex experimentation without peril, confident that the simulated environment poses no danger to human operators or expensive equipment. This enhanced safety is pivotal, particularly in domains where failure is not an option.

Key Methodologies

Sim2Real encompasses a range of methodologies, each tailored to specific applications and domains. A fundamental aspect in this pursuit is achieving high-fidelity simulations that closely mimic real-world conditions. To narrow the gap between the data distributions of simulated environments and actual real-world scenarios, several techniques have emerged as key contributors.

Domain randomization, a widely used technique in the realm of ML and Sim2Real transfer, plays a crucial role in enhancing the adaptability and robustness of models [8, 9]. This technique involves training models across a variety of simulated environments, each distinguished by randomized characteristics. The objective is to instill ML models with the capability to effectively manage uncertainty and adapt to unforeseen variations, commonly encountered in real-world settings.

Numerous studies have explored the effectiveness of Sim2Real transfer methods based on domain randomization [10–14]. These studies have demonstrated the potential of domain randomization in creating a diverse and extensive training dataset. Figure 1, which illustrates several variations of low-fidelity training images with random camera

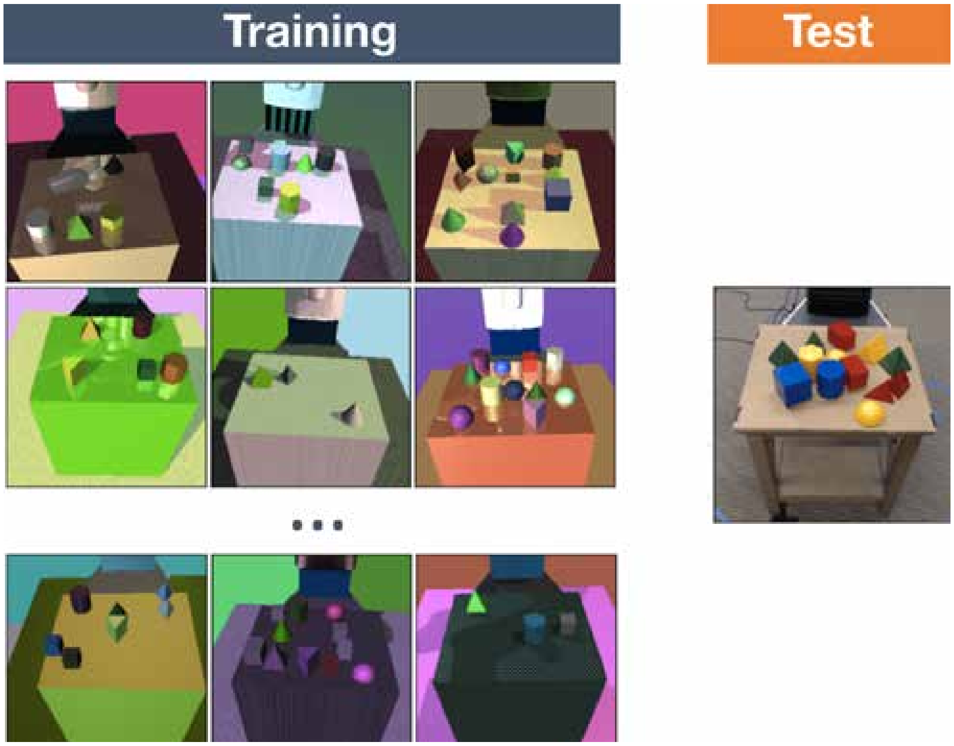


Figure 1. Variations of Low-Fidelity Training Images for Domain Randomization *(Source: Tobin et al. [10]).*

positions, lighting conditions, object positions, and nonrealistic textures, showcases the application of domain randomization in generating a robust training set. This diversity enables models to excel in making accurate predictions when confronted with the intricacies of real-world environments, even when faced with previously unseen conditions during training.

Adversarial training, a widely adopted technique in deep learning (DL), focuses on enhancing the robustness and security of ML models [15–18]. It introduces adversarial examples during training, which, while often imperceptible to humans, perturb input data to deliberately induce incorrect predictions from the model. Including adversarial examples in the training data renders the model less susceptible to manipulation and significantly improves its performance in terms of robustness in the presence of noise and adversarial inputs.

Adversarial training plays a pivotal role in addressing the Sim2Real transfer problem, where models trained in the controlled environments of simulations are required to perform seamlessly in unpredictable real-world conditions [19–22]. Recent research underlines the critical role of adversarial training in minimizing domain discrepancies and enhancing model adaptability. By diminishing the gap between the realm of simulation and that of reality, this technique offers substantial value across a wide array of applications, ranging from robotics to autonomous systems.

CURRENT APPLICATIONS OF SIM2REAL

Sim2Real techniques represent a significant innovation within the ML domain. Their transformative potential has been most prominently realized in robotics and autonomous driving [3], where they have been rigorously tested and refined. In this section, the concrete applications of Sim2Real within these domains are explored, shedding light on their impact and efficacy in addressing real-world challenges.

The core applications of Sim2Real in robotics are examined first, where simulated environments prove to be heavily effective for training and optimizing intelligent systems. Sim2Real's ability to bridge the gap between simulation and reality has empowered robots to interact seamlessly with their surroundings, enabling the navigation of complex terrains [23–26] and object detection, recognition, and manipulation with precision [11, 27–30].

Similarly, in the domain of autonomous driving, Sim2Real techniques have played a pivotal role in enhancing vehicle autonomy and safety. By leveraging simulated environments, autonomous vehicles have undergone extensive training, enabling them to navigate diverse road conditions [31–35] and respond to complex scenarios [35–38].

As the applications of Sim2Real in these well-established domains are traversed, the broader horizons are explored, where these techniques have the potential to reshape and revolutionize ML applications across various fields, including the cyber domain.

Applications of Sim2Real in Robotics

Sim2Real techniques have made significant strides in the realm of robotics, reshaping the landscape of intelligent systems' capabilities and adaptability. These approaches seamlessly bridge the gap between simulated environments and real-world applications, equipping robots with a diverse range of capabilities.

Navigation of Complex Terrains

In the quest to navigate intricate and challenging terrains, simulated environments have become crucial training grounds. Deep reinforcement learning (DRL) plays a pivotal role in these advancements, enabling robots to adapt and excel in real-world but simulated scenarios.

As introduced in Hu et al. [23], a novel Sim2Real pipeline empowers mobile robots to navigate three-dimensional (3-D) rough terrains. The pipeline not only facilitates successful

> *In the quest to navigate intricate and challenging terrains, simulated environments have become crucial training grounds.*

point-to-point navigation but also outperforms classical and other DL-based approaches in terms of success rate, cumulative travel distance, and time. Comprehensive surveys, such as the one in Zhao, Queralta, and Westerlund [24], shed light on the challenges of Sim2Real transfer in DRL, categorizing various approaches aimed at closing the gap between simulated and real-world performance, especially in navigation.

Moreover, Figure 2 demonstrates a hybrid architecture [25] that effectively employs attention-based DRL for navigation cost map generation in outdoor environments. This approach addresses the challenge of obstacle avoidance in complex terrains, as discussed in Zhang et al. [26]. By following least-cost waypoints on the cost map, the robot significantly enhances its performance in uneven outdoor terrains. These examples highlight the versatility of Sim2Real techniques in addressing complex terrains, making robots more adaptable and efficient in navigating challenging landscapes.

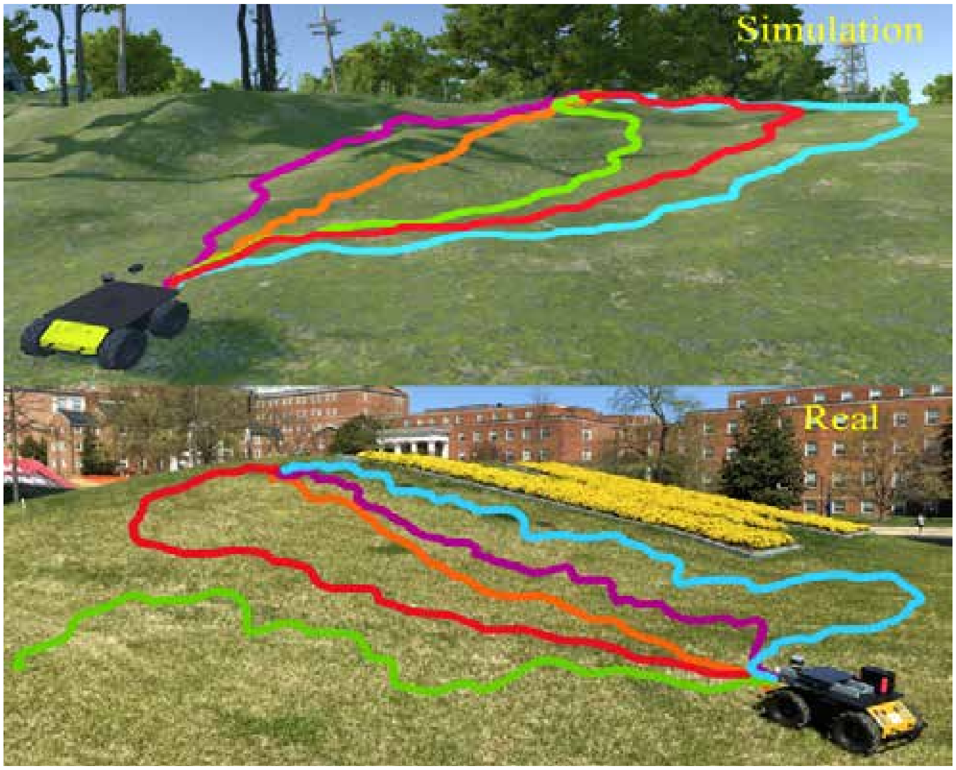


Figure 2. Comparison of Navigation Methods on Uneven Outdoor Terrains *(Source: Weerakoon et al. [25])*.

Object Detection, Recognition, and Manipulation

Precision in object detection, recognition, and manipulation is a hallmark of advanced robotics. Researchers in Ho et al. [27] introduced innovative techniques such as RetinaGAN to address the challenge of collecting real-world data for training DRL and imitation learning ML models. This generative adversarial network approach adapts simulated images to resemble real-world scenes with object-detection consistency. The result is a substantial improvement in the performance of reinforcement learning-based tasks like object instance grasping, pushing, and even the more complex task of door opening.

Sim2Real techniques have also been instrumental in object detection, with domain randomization serving as a key method, as highlighted in Horváth et al. [11]. By generating labeled synthetic datasets at scale, Sim2Real transfer learning ensures that state-of-the-art convolutional neural networks, such as YOLOv4, can achieve impressive mean average precision scores in scenarios where labeled real-world data may be scarce. Furthermore, in the industrial sector, Sim2Real techniques enable fast and accurate object recognition and localization for robotic bin picking, as detailed in Li et al. [28]. Supported by automated synthetic data generation pipelines, these methods not only provide precise training data but also excel in scenarios involving textureless, metallic, and occluded objects.

The application of Sim2Real extends to 3-D object detection from point clouds, a domain known for its challenges, as illustrated in DeBortoli et al. [29]. By leveraging adaptive sampling modules and 3-D adversarial training architectures, Sim2Real approaches enhance the consistency of features extracted from point clouds, improving 3-D object detection performance.

Even deformable objects like cloth are not beyond the reach of Sim2Real methods. Deformable object manipulation is a relatively unexplored frontier, with a notable data shortfall. In Matas et al. [30], agents are trained entirely in simulation, using domain randomization to ensure their versatility. They are then successfully deployed in the real world without prior exposure to real deformable objects.

In summary, Sim2Real techniques have evolved into indispensable tools, empowering robots to navigate complex terrains and execute precise object detection, recognition, and manipulation tasks. These advancements have created an era of highly adaptable and capable robotic systems, setting the stage for a new wave of innovations in intelligent robotics.

As Sim2Real in autonomous driving is explored, the notion that Sim2Real techniques can extend not only within the boundaries of robotics but into the broader space of complex real-world challenges is proposed.

Applications of Sim2Real for Autonomous Driving

Autonomous driving is a complex field that demands intelligent vehicles capable of navigating diverse road conditions and responding to complex scenarios. Integrating Sim2Real techniques has played a pivotal role in enhancing vehicle autonomy and safety by bridging the gap between simulated training environments and real-world deployment. How Sim2Real techniques enable autonomous vehicles to excel in challenging conditions is explored next.

Navigating Diverse Road Conditions

In the realm of autonomous driving, it is imperative that vehicles can navigate diverse road conditions, from off-road terrains to urban environments. Sim2Real techniques offer innovative solutions to train and deploy autonomous vehicles effectively.

A human-guided RL framework is introduced in Wu et al. [31], enhancing the learning process and capabilities of RL methods. This approach allows humans to intervene in the control progress, providing demonstrations as needed. The result is a versatile RL agent trained in simulation and effectively transferred to real-world unmanned ground vehicles, demonstrating robust navigation in dynamic and diverse environments.

> “
>
> ***Sim2Real techniques have evolved into indispensable tools, empowering robots to navigate complex terrains and execute precise object detection, recognition, and manipulation tasks.***

To further bridge the visual reality gap for off-road autonomous driving, So et al. [32] introduce Sim2Seg, a novel approach that translates randomized simulation images into simulated segmentation and depth maps, enabling the direct deployment of an end-to-end RL policy in real-world scenarios. Sim2Seg effectively narrows the gap between the simulated training environment and real-world driving, particularly in off-road conditions.

An approach that combines the advantages of modular architectures and end-to-end DL for autonomous driving is presented by Müller et. al [33]. By encapsulating the driving policy, it successfully transfers policies trained in simulation to real-world deployments, addressing the challenges of adapting to diverse road conditions.

Mapless navigation is a crucial aspect of autonomous driving, and Wang et al. [34] propose a DRL-based approach for unmanned surface vehicles. This method carefully designs observation and action spaces and rewards functions and neural networks for navigation policies. By employing domain randomization and adaptive curriculum learning, it offers an effective solution to the Sim2Real transfer challenge and slow convergence associated with DRL.

A data-driven simulation and training engine, which allows autonomous vehicles to learn end-to-end control policies through sparse rewards, is discussed in Amini et al. [35]. As illustrated in Figure 3, training images from several comparison methods used in experimentation highlight the diversity of environments and conditions encountered during training. This simulation enables vehicles to navigate a continuum of new local trajectories in diverse road conditions, proving the feasibility of transferring policies from simulation to real-world deployment. It also helps transition into the discussion on Sim2Real's ability to help with the response to complex scenarios.

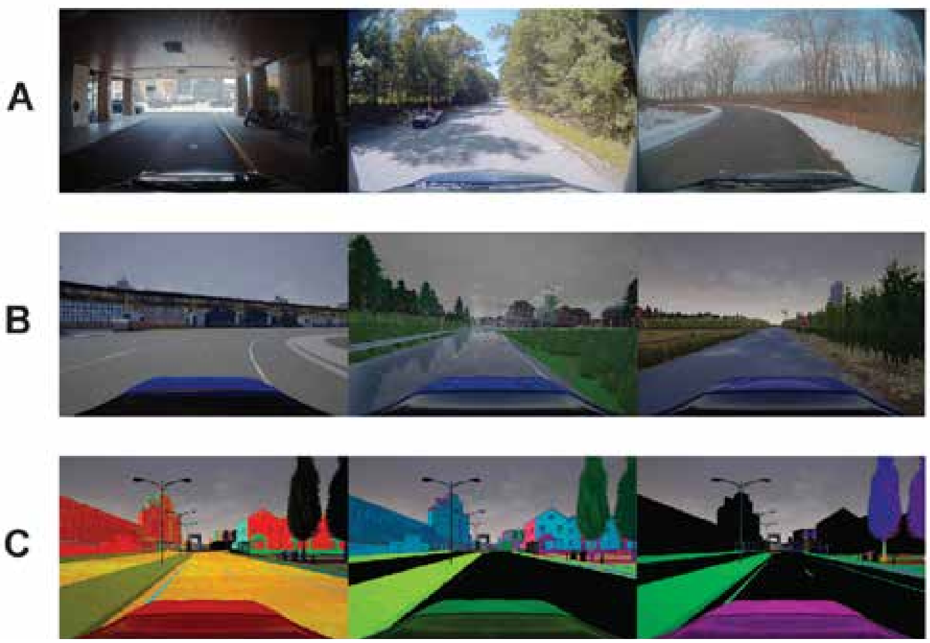


Figure 3. Training Images From (A) Real-World Samples and (B–C) Simulated Environments *(Source: Amini et al. [35]).*

Responding to Complex Scenarios

Autonomous vehicles must not only navigate diverse road conditions but also respond effectively to complex scenarios, including near-crash situations and challenging traffic interactions. How Sim2Real techniques equip these vehicles to tackle intricate real-world challenges is analyzed.

In the context of responding to complex scenarios, Amini et al. [35] present a data-driven simulation and training engine that learns end-to-end autonomous vehicle control policies using sparse rewards. By rendering novel training data derived from real-world trajectories, the simulator allows virtual agents to navigate previously unseen real-world roads, even in near-crash scenarios. This approach demonstrates the potential of Sim2Real techniques to create policies that can handle complex and novel situations.

A method that transfers a vision-based lane following driving policy from simulation to real-world operation on rural roads without any real-world labels is introduced in Bewley et al. [36]. Leveraging image-to-image translation, a single-camera control policy is learned while achieving domain transfer. This approach successfully operates autonomous vehicles in rural and urban environments, illustrating the applicability of Sim2Real for complex real-world scenarios.

Unsupervised domain adaptation methods have been developed for lane detection and classification in autonomous driving, as described in Hu et al. [37]. By using synthetic data generated in simulation, these methods leverage adversarial discriminative and generative techniques to adapt to the real world. They demonstrate superiority in detection and classification accuracy and consistency in complex traffic scenarios.

Complex multivehicle and multilane scenarios are particularly challenging for autonomous vehicles. A Sim2Real approach to safely learn driving policies for autonomous vehicles sharing the road with other vehicles and obstacles is discussed in Mitchell et al. [38]. This approach leverages mixed reality setups to simulate collisions and interactions, making the learning process safer. After only a few runs in mixed reality, collisions are significantly reduced, indicating the approach's effectiveness in addressing complex traffic scenarios.

Sim2Real's influence within the autonomous driving domain has transcended conventional boundaries, echoing its success in robotics. These techniques have introduced advancements in self-driving technology, enhancing not only the capabilities but also the safety of autonomous vehicles.

As the potential applications of Sim2Real within the cyber domain are explored next, transforming Sim2Real is far from over. Figure 4 portrays the intricate overlap between these existing and potential applications, offering a nuanced perspective on the transformative potential that Sim2Real introduces across diverse domains. Just as it has reshaped the realms of robotics and autonomous driving, Sim2Real now holds the promise of innovative solutions and novel perspectives to address the multifaceted complexities of the cyber domain.

POTENTIAL APPLICATIONS OF SIM2REAL IN THE CYBER DOMAIN

In the cyber domain, advancement is not merely a desire but a necessity. However, when it comes to applying Sim2Real techniques, a noticeable void remains. Unlike its well-established presence in robotics and autonomous driving, Sim2Real remains largely unexplored within the cyber domain. This limitation can be attributed to the nascent application of ML to cybersecurity and the constrained availability of advanced cyber simulators and emulators.

This absence prompts a pivotal question: What can the cyber domain gain from Sim2Real techniques? The answer lies in the intrinsic nature of cybersecurity—a realm where the stakes are high and the consequences of failure can be catastrophic. In an era where cyber threats perpetually evolve, organizations require dynamic, adaptable, and data-rich environments for training, testing, and fortifying their defenses. Sim2Real holds the potential to bridge this gap.

The opportunities that this fusion of simulation and reality can unlock, from reshaping cybersecurity training and testing to facilitating meticulous vulnerability assessments, are discussed next. As exploration begins, Sim2Real's potential is offered—an opportunity to redefine the landscape of the cyber domain in a way that is not only innovative but indispensable.

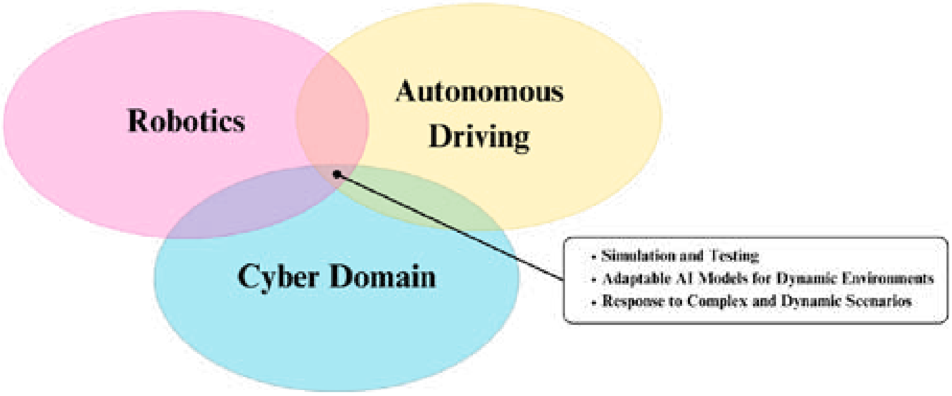

Robotics
Autonomous Driving
Cyber Domain
• Simulation and Testing
• Adaptable AI Models for Dynamic Environments
• Response to Complex and Dynamic Scenarios

Figure 4. Visualizing the Overlap Between Existing and Potential Applications of Sim2Real *(Source: E. Nack).*

Cybersecurity Training and Testing

Various approaches have been employed in cybersecurity training and testing to equip professionals with the necessary skills and experience. Traditional classroom-style training provides theoretical instruction to individuals entering cybersecurity [39, 40]. However, this approach often falls short in replicating the real-world dynamics and pressures associated with cyber threats.

To enhance engagement and provide a more immersive experience, the industry has turned to gamification. Several training platforms incorporate gamified scenarios, allowing participants to navigate simulated cyber threats interactively [41, 42]. "Capture The Flag" competitions are a compelling example of gamification within cybersecurity, providing a platform for participants to solve security-related puzzles and challenges [43]. Although widely recognized for its positive impact on user engagement and skill development, gamification often focuses on specific elements of cybersecurity, providing expertise in targeted areas but not offering a comprehensive training method for the diverse and rapidly evolving cyber threat landscape.

These current practices, while contributing significantly to cybersecurity training and testing, face inherent challenges. The limitations of predefined scenarios, the static nature of exercises, and the finite set of challenges in gamified platforms all point to the need for dynamic and adaptable training environments capable of replicating the intricacies of real-world cyber threats.

> ***Although widely recognized for its positive impact on user engagement and skill development, gamification often focuses on specific elements of cybersecurity.***

Recognizing the limitations in current practices, the integration of Sim2Real techniques becomes a natural progression. While existing simulations provide valuable training and testing environments [39, 44, 45], the critical aspect of seamlessly transitioning from simulation to reality remains underexplored. The literature on Sim2Real techniques in cybersecurity is limited, particularly concerning the effectiveness of applying knowledge gained in simulated environments to real-world cybersecurity scenarios.

In this context, Cyber Virtual Assured Network (CyberVAN) and Cyber Battlefield Operating System Simulation (CyberBOSS) emerge as valuable simulation environments that address the existing gaps in cybersecurity training and testing. CyberVAN functions as a discrete event simulator, offering a quick and flexible setup of high-fidelity cybersecurity scenarios [46]. This simulation environment provides a dynamic and realistic platform for training and testing, enabling the instantiation of custom high-fidelity networks. Its capability aligns seamlessly with the customization potential of Sim2Real environments to mirror specific network architectures, industry sectors, or regulatory compliance requirements, enhancing the overall adaptability of the training environment. Furthermore, as Sim2Real techniques aim to open doors to dynamic and realistic cyber environments, CyberVAN contributes by replicating actual cyber threats and scenarios in controlled settings. Serving as crucial training grounds, these simulations empower cybersecurity professionals to refine their detection, mitigation, and response strategies effectively.

In addition to CyberVAN, CyberBOSS introduces a framework designed to model cyberspace effects and operations [47]. This framework operates across federated live, virtual, constructive, and gaming (LVC&G) systems, offering a comprehensive solution for cyberspace modeling in environments that may lack these capabilities. CyberBOSS stands as a strategic asset, particularly in scenarios where native cyberspace modeling is limited or nonexistent. Figure 5

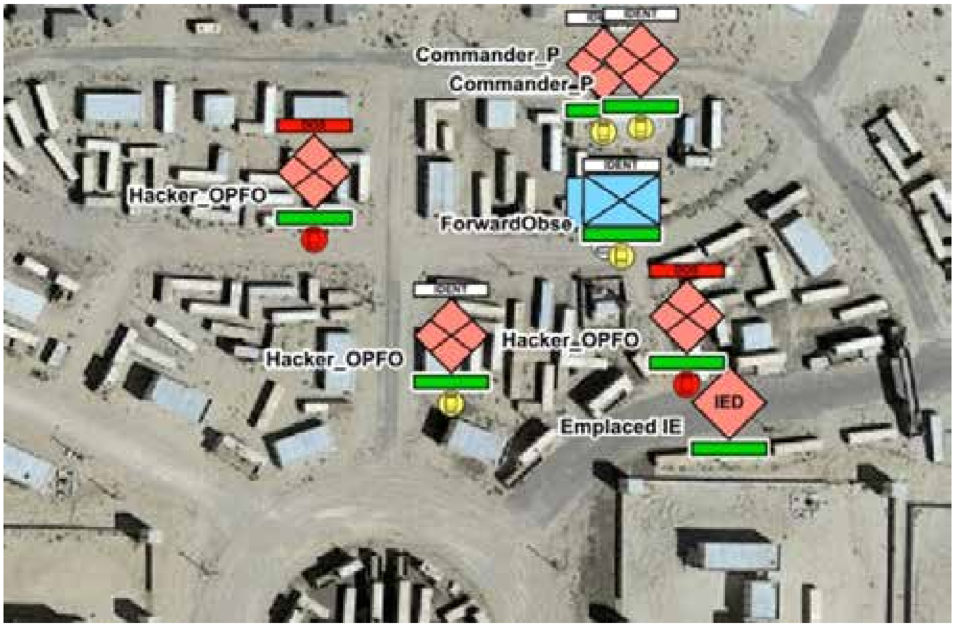


Figure 5. Visualizing Cyberspace-Related Objects and Effects Using CyberBOSS *(Source: Hasan et al. [47]).*

visualizes an example of cyberspace-related objects and effects within the STTC's Battlespace Visualization and Interaction tool using information provided by the CyberBOSS federation. Integrating these advanced simulation tools, combined with Sim2Real techniques, enhances the potential of bridging the gap between simulation training scenarios and the complex realities of cybersecurity, ensuring that professionals remain agile and well-prepared against emerging threats.

In conclusion, the current research gap in Sim2Real within cybersecurity necessitates further exploration and investigation into the seamless transition from simulated cybersecurity environments to real-world applications; however, integrating advanced simulation tools like CyberVAN and CyberBOSS stands as a promising avenue for addressing this gap and revolutionizing cybersecurity training and testing.

Vulnerability Testing

Vulnerability testing stands as a cornerstone of effective cybersecurity, playing a pivotal role in the timely identification and mitigation of potential threats to an organization's digital assets. While established methods like vulnerability assessment and penetration testing have made significant contributions to this process, inherent limitations prompt a reevaluation of these existing approaches.

Vulnerability assessment stands as a passive and proactive strategy, employing tools like Nessus [48] and OpenVAS [49] to systematically uncover known vulnerabilities in network configurations and software systems. A comparative study of these tools highlights their features and effectiveness in vulnerability assessment [50]. However, as automation proliferates, organizations grapple with copious data, including false positives. Challenges also extend to identifying logical attack vectors, such as application logic flaws and password reuse, often resulting in generic remediation recommendations. The rigidity of automated tools may also lead to oversight in emerging threats or vulnerabilities not present in their databases [51]. This limitation hampers effectiveness in adapting to the dynamically evolving threat landscapes.

On the other hand, penetration testing embraces an active and systematic approach to security assessment. Ethical hacking simulates real-world cyber attacks, providing profound insights into the impact of identified vulnerabilities on the information system [52]. This exploration considers mitigating controls and allows for a comprehensive evaluation of the security landscape, eliminating false positives. Yet, penetration testing demands considerable time and effort, potentially requiring external engagement for comprehensive testing. Its outcomes may not guarantee identifying every vulnerability, and it might not provide insights into emerging vulnerabilities post assessment [51].

Confronting these challenges embedded in traditional vulnerability assessment and penetration testing, Sim2Real techniques emerge as a transformative solution. By allowing testing and assessment to take place off the network, Sim2Real provides a controlled and secure environment for organizations to simulate various cyber attacks and exploitation techniques, gaining insights before deployment into the live network. This proactive approach empowers organizations to identify and address vulnerabilities before they can be exploited by malicious actors.

Furthermore, Sim2Real assessments extend the scope of vulnerability testing beyond traditional methods, accommodating complex infrastructures, including cloud-based systems, Internet of Things (IoT) devices, and critical infrastructure networks—all within a secure and controlled context.

In essence, Sim2Real techniques offer a unique set of advantages, enabling organizations to map and analyze more complex scenarios before deployment and ultimately enhancing their cybersecurity posture.

Training AI for Cybersecurity

> “
>
> ***As automation proliferates, organizations grapple with copious data, including false positives.***

Artificial intelligence (AI) is rapidly evolving in modern cybersecurity, empowering threat detection, anomaly detection, predictive analysis, and more. However, training AI models for cybersecurity applications demands large amounts of diverse and high-quality data [53]. Currently, researchers rely on popular datasets such as ACI-IoT-2023 [54], KDD’99 Cup [55], UNSW-NB15 [56], CIC-IDS2017 [57], CIC-DDoS2019 [58], CERT [59, 60], and Bot-IoT [61]. These datasets encompass a range of cyber threats, providing a foundation for training AI models to recognize and respond to various attack vectors, vulnerabilities, and malicious activities.

However, building comprehensive datasets for AI training in cybersecurity poses significant challenges. Acquiring diverse and realistic data often requires a well-equipped cybersecurity lab setup, a multitude of devices, and precise data collection methods. Constructing such datasets is not only resource-intensive but can also be challenging due to the dynamic nature of cyber threats. The complexities involved in replicating real-world cyber scenarios highlight the limitations of current approaches in providing sufficiently diverse and adaptive datasets for effective AI model training.

Sim2Real techniques provide a compelling and more effective solution for training AI models by creating data-rich environments that accurately mimic the complexities of actual cyber landscapes. This novel approach addresses the limitations of existing training practices, offering a dynamic and realistic environment for AI models to learn and adapt to the ever-evolving landscape of cyber threats.

Within Sim2Real-based cyber environments, AI models can be exposed to a vast array of realistic cyber threats and scenarios. These environments generate diverse and dynamic datasets that encompass various attack vectors, malware samples, and network traffic patterns. AI models can learn from these simulated encounters, improving their ability to detect and respond to real-world cyber threats effectively.

Moreover, Sim2Real allows for the injection of controlled anomalies and variations into the data, enabling AI models to develop robust anomaly detection capabilities. AI models trained in these environments become highly adaptable, as they are exposed to a broad spectrum of cyber scenarios from routine network traffic to sophisticated zero-day attacks.

The implications of training AI for cybersecurity within Sim2Real environments extend to predictive analysis and threat intelligence. AI models can learn to recognize patterns indicative of emerging threats, enhancing an organization's ability to proactively respond to evolving cyber risks.

CHALLENGES AND LIMITATIONS

Despite the promising potential of Sim2Real techniques in the cyber domain, their implementation faces certain challenges and limitations that should be considered.

Simulations, while valuable, often encounter a significant hurdle known as the "Simulation to Reality Gap" [62, 63]. These simulations can replicate real-world scenarios to a certain extent, but they may not fully capture all complexities and the unpredictable nature. This inherent discrepancy could potentially impact the effectiveness of Sim2Real techniques in preparing for and responding to real-world cyber threats.

Data privacy and security are significant concerns in any application that involves extensive data [64]. While simulations often demand extensive data for efficacy, incorporating real-world data into simulations introduces potential privacy risks, especially within the cyber domain where sensitive information is often involved. Striking a delicate balance between ensuring data privacy and maintaining simulation effectiveness becomes crucial in this context.

Furthermore, the quantity and quality of data are critical considerations in Sim2Real techniques. Insufficient or inaccurate data can impede the fidelity of simulations, hindering their effectiveness in modeling real-world scenarios. Recognizing and addressing these challenges is essential in enhancing the accuracy and reliability of these techniques.

A fundamental challenge lies in validating Sim2Real techniques within the cyber domain. Given the relatively unexplored nature of Sim2Real in this domain, the absence of established metrics or benchmarks hampers evaluating the success of these techniques. For instance, without standardized measures, it becomes challenging to assess how well a model trained in a simulated environment would perform when facing real-world cyber threats.

Despite these challenges, recognizing the potential benefits of Sim2Real techniques in the cyber domain is crucial. Continued research and development in this field hold the promise of revolutionizing cybersecurity training and testing, offering a dynamic and adaptable solution to the ever-evolving landscape of cyber threats. However, it is essential to approach this potential with a clear understanding of the challenges and limitations, ensuring realistic expectations and effective strategies for overcoming these obstacles.

> ***While simulations often demand extensive data for efficacy, incorporating real-world data into simulations introduces potential privacy risks.***

CONCLUSIONS

The convergence of Sim2Real techniques with the cyber domain represents a promising frontier in cybersecurity and ML. This article delved into the foundational concepts of Sim2Real, explored its current applications in robotics and autonomous driving, and examined its potential applications, challenges, and limitations within the cyber domain.

While Sim2Real techniques have already demonstrated their value in training robots and autonomous vehicles, their application within the cyber domain holds the potential to revolutionize how organizations prepare for, defend against, and respond to cyber threats. Creating dynamic and data-rich cyber environments for training, testing, vulnerability assessment, and AI model training offers a new paradigm for enhancing cybersecurity resilience.

Despite the immense promise, integrating Sim2Real techniques into the cyber domain is not without its challenges. However, these challenges present compelling research opportunities for the future of Sim2Real in the cyber domain. Exploring the seamless integration of Sim2Real techniques with real-world cybersecurity scenarios, optimizing adaptability in the face of evolving threats, establishing standardized metrics for assessment, and refining the transferability of simulated knowledge to practical applications are just a few examples of the exciting avenues for future exploration.

In the future, Sim2Real gaps will be evaluated by training an ML-based Network Intrusion Detection System classifier in CyberVAN. Subsequently, the objective is to transfer this classifier to the IoT Research Lab at the Army Cyber Institute to assess its effectiveness and explore ways to improve performance. Such experimentation will contribute significantly to understanding the challenges and potentials of Sim2Real techniques in real-world cyber applications.

> ***Creating dynamic and data-rich cyber environments for training, testing, vulnerability assessment, and AI model training offers a new paradigm for enhancing cybersecurity resilience.***

By embracing these techniques and addressing their challenges, the cyber domain can become more adaptive, resilient, and prepared to safeguard the digital ecosystems that underpin modern society. Continued research and development will be key to unlocking the full potential of Sim2Real in enhancing cyber resilience and redefining the landscape of cybersecurity practices. ■

REFERENCES

[1] Hu, X., et al. "How Simulation Helps Autonomous Driving: A Survey of Sim2real, Digital Twins, and Parallel Intelligence." *IEEE Transactions on Intelligent Vehicles*, 2023.

[2] Höfer, S., et al. "Perspectives on Sim2real Transfer for Robotics: A Summary of the R: Ss 2020 Workshop." arXiv preprint arXiv:2012.03806, 2020.

[3] Höfer, S., et al. "Sim2Real in Robotics and Automation: Applications and Challenges." *IEEE Transactions on Automation Science and Engineering*, vol. 18.2, pp. 398–400, 2021.

[4] Salvato, E., et al. "Crossing the Reality Gap: A Survey on Sim-To-Real Transferability of Robot Controllers in Reinforcement Learning." *IEEE Access*, vol. 9, pp. 153171–153187, 2021.

[5] Chebotar, Y., et al. "Closing the Sim-to-Real Loop: Adapting Simulation Randomization With Real World Experience." The 2019 International Conference on Robotics and Automation (ICRA), IEEE, 2019.

[6] Tan, J., et al. "Sim-to-Real: Learning Agile Locomotion for Quadruped Robots." arXiv preprint arXiv:1804.10332, 2018.

[7] James, S., et al. "Sim-to-Real via Sim-to-Sim: Data-Efficient Robotic Grasping via Randomized-to-Canonical Adaptation Networks." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.

[8] Chen, X., et al. "Understanding Domain Randomization for Sim-to-Real Transfer." arXiv preprint arXiv:2110.03239, 2021.

[9] Muratore, F., et al. "Robot Learning From Randomized Simulations: A Review." *Frontiers in Robotics and AI*, vol. 31, 2022.

[10] Tobin, J., et al. "Domain Randomization for Transferring Deep Neural Networks From Simulation to the Real World." The 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2017.

[11] Horváth, D., et al. "Object Detection Using Sim2real Domain Randomization for Robotic Applications." *IEEE Transactions on Robotics*, vol. 39.2, pp. 1225–1243, 2022.

[12] Marez, D., S. Borden, and L. Nans. "UAV Detection With a Dataset Augmented by Domain Randomization." *Geospatial Informatics X*, vol. 11398, SPIE, 2020.

[13] Ranaweera, M., and Q. Mahmoud. "Evaluation of Techniques for Sim2Real Reinforcement Learning." *The International FLAIRS Conference Proceedings*, vol. 36, 2023.

[14] Akkaya, I., et al. "Solving Rubik's Cube With a Robot Hand." arXiv preprint arXiv:1910.07113, 2019.

[15] Kurakin, A., I. Goodfellow, and S. Bengio. "Adversarial Machine Learning at Scale." The 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings, arXiv preprint arXiv:1611.01236, November 2016.

[16] Huang, J., H. J. Choi, and N. Figueroa. "Trade-Off Between Robustness and Rewards Adversarial Training for Deep Reinforcement Learning Under Large Perturbations." IEEE Robotics and Automation Letters, 2023.

[17] Shrivastava, A., et al. "Learning From Simulated and Unsupervised Images Through Adversarial Training." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[18] Tramèr, F., A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel. "Ensemble Adversarial Training: Attacks and Defenses." In the International Conference on Learning Representations (ICLR), arXiv preprint arXiv:1705.07204, https://arxiv.org/abs/1705.07204, 2017.

[19] Rezaeianaran, F., et al. "Seeking Similarities over Differences: Similarity-Based Domain Alignment for Adaptive Object Detection." Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.

[20] Zhai, P., et al. "Robust Adaptive Ensemble Adversary Reinforcement Learning." *IEEE Robotics and Automation Letters*, vol. 7.4, pp. 12562–12568, 2022.

[21] Shin, I., K. Park, S. Woo, and I. S. Kweon. "Unsupervised Domain Adaptation for Video Semantic Segmentation." In the AAAI Conference on Artificial Intelligence, arXiv preprint arXiv:2107. 11052, 2022.

[22] Cheng, K., C. Healey, and T. Wu. "Towards Adversarially Robust and Domain Generalizable Stereo Matching by Rethinking DNN Feature Backbones." arXiv preprint arXiv:2108.00335, 2021.

[23] Hu, H., et al. "A Sim-to-Real Pipeline for Deep Reinforcement Learning for Autonomous Robot Navigation in Cluttered Rough Terrain." *IEEE Robotics and Automation Letters*, vol. 6.4, pp. 6569–6576, 2021.

[24] Zhao, W., J. P. Queralta, and T. Westerlund. "Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: A Survey." 2020 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2020.

[25] Weerakoon, K., A. J. Sathyamoorthy, and D. Manocha. "Sim-to-Real Strategy for Spatially Aware Robot Navigation in Uneven Outdoor Environments." In the ICRA 2022 Workshop on Releasing Robots Into the Wild, arXiv preprint arXiv:2205.09194, 2022.

[26] Zhang, T., et al. "Sim2real Learning of Obstacle Avoidance for Robotic Manipulators in Uncertain Environments." *IEEE Robotics and Automation Letters*, vol. 7.1, pp. 65–72, 2021.

[27] Ho, D., et al. "Retinagan: An Object-Aware Approach to Sim-to-Real Transfer." The 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021.

[28] Li, X., et al. "A Sim-to-Real Object Recognition and Localization Framework for Industrial Robotic Bin Picking." *IEEE Robotics and Automation Letters*, vol. 7.2, pp. 3961–3968, 2022.

[29] DeBortoli, R., et al. "Adversarial Training on Point Clouds for Sim-to-Real 3D Object Detection." IEEE Robotics and Automation Letters 6.4, pp. 6662–6669, 2021.

[30] Matas, J., S. James, and A. J. Davison. "Sim-to-Real Reinforcement Learning for Deformable Object Manipulation." Conference on Robot Learning, PMLR, 2018.

[31] Wu, J., et al. "Human-Guided Reinforcement Learning With Sim-to-Real Transfer for Autonomous Navigation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[32] So, J., A. Xie, S. Jung, J. Edlund, R. Thakker, A. Agha-Mohammadi, P. Abbeel, and S. James. "Sim-to-Real via Sim-to-Seg: End-to-End Off-Road Autonomous Driving Without Real Data." Proceedings of the 6th Conference on Robot Learning, PMLR 205, pp. 1871–1881, arXiv preprint arXiv: 2210.14721, 2022.

[33] Müller, M., A. Dosovitskiy, B. Ghanem, and V. Koltun. "Driving Policy Transfer via Modularity and Abstraction." Proceedings of the 2nd Conference on Robot Learning in *Proceedings of Machine Learning Research*, vol. 87, pp. 1–15, https://proceedings.mlr.press/v87/mueller18a.html, arXiv preprint arXiv:1804.09364, 2018.

[34] Wang, N., et al. "Sim-to-Real: Mapless Navigation for USVs Using Deep Reinforcement Learning." *Journal of Marine Science and Engineering*, vol. 10.7, p. 895, 2022.

[35] Amini, A., et al. "Learning Robust Control Policies for End-to-End Autonomous Driving From Data-Driven Simulation." *IEEE Robotics and Automation Letters*, vol. 5.2, pp. 1143–1150, 2020.

[36] Bewley, A., et al. "Learning to Drive From Simulation Without Real World Labels." The 2019 International Conference on Robotics and Automation (ICRA), IEEE, 2019.

[37] Hu, C., et al. "Sim-to-Real Domain Adaptation for Lane Detection and Classification in Autonomous Driving." The 2022 IEEE Intelligent Vehicles Symposium (IV), 2022.

[38] Mitchell, R., J. Fletcher, J. Panerati, and A. Prorok. "Multi-Vehicle Mixed Reality Reinforcement Learning for Autonomous Multi-Lane Driving." In the Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, Auckland, New Zealand, pp. 1928–1930, arXiv preprint arXiv:1911.11699, 2020.

[39] Urias, V. E., et al. "Dynamic Cybersecurity Training Environments for an Evolving Cyber Workforce." The 2017 IEEE International Symposium on Technologies for Homeland Security (HST), IEEE, 2017.

[40] Hatzivasilis, G., et al. "Modern Aspects of Cyber-Security Training and Continuous Adaptation of Programmes to Trainees." *Applied Sciences*, vol. 10.16, p. 5702, 2020.

[41] Jin, G., et al. "Evaluation of Game-Based Learning in Cybersecurity Education for High School Students." *Journal of Education and Learning (EduLearn)*, vol. 12.1, pp. 150–158, 2018.

[42] Jin, G., et al. "Game Based Cybersecurity Training for High School Students." Proceedings of the 49th ACM Technical Symposium on Computer Science Education, 2018.

[43] Švábenský, V., et al. "Enhancing Cybersecurity Skills by Creating Serious Games." Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education, 2018.

[44] Chowdhury, N., and V. Gkioulos. "Cyber Security Training for Critical Infrastructure Protection: A Literature Review." *Computer Science Review*, vol. 40, p. 100361, 2021.

[45] Prümmer, J., T. van Steen, and B. van den Berg. "A Systematic Review of Current Cybersecurity Training Methods." *Computers & Security*, p. 103585, 2023.

[46] Chadha, R., et al. "Cybervan: A Cyber Security Virtual Assured Network Testbed." MILCOM 2016-IEEE Military Communications Conference, 2016.

[47] Hasan, O., J. Geddes, J. Welch, N. Vey, and R. Burch. "A Cyberspace Effects Server for LVC&G Training Systems." I/ITSEC 2021 Conference (paper no. 21258), Orlando, FL, 2021.

[48] Tenable. "Nessus Vulnerability Scanner: Network Security Solution." www.tenable.com/products/nessus, accessed 12 December 2023.

[49] OpenVAS – Open Vulnerability Assessment Scanner. www.openvas.org, accessed 12 December 2023.

[50] Yadav, S. K., D. S. Pandey, and S. Lade. "A Comparative Analysis of Detecting Vulnerability in Network Systems." *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 7.5, 2017.

[51] Shinde, P. S., and S. B. Ardhapurkar. "Cyber Security Analysis Using Vulnerability Assessment and Penetration Testing." The 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave), IEEE, 2016.

[52] Yaqoob, I., et al. "Penetration Testing and Vulnerability Assessment." *Journal of Network Communications and Emerging Technologies (JNCET)*, vol. 7.8, www.jncet.org, 2017.

[53] Sarker, I. H., et al. "Cybersecurity Data Science: An Overview from Machine Learning Perspective." *Journal of Big Data*, vol. 7, pp. 1–29, 2020.

[54] Bastian, N., D. Bierbrauer, M. McKenzie, and E. Nack. "ACI IoT Network Traffic Dataset 2023." IEEE Dataport, https://dx.doi.org/10.21227/qacj-3x32, 29 December 2023.

[55] Tavallaee, M., et al. "A Detailed Analysis of the KDD CUP 99 Data Set." The 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009.

[56] Moustafa, N., and J. Slay. "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set)." The 2015 Military Communications and Information Systems Conference (MilCIS), IEEE, 2015.

[57] Sharafaldin, I., A. H. Lashkari, and A. A. Ghorbani. "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization." ICISSP 1, pp. 108-116, 2018.

[58] Sharafaldin, I., et al. "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy." The 2019 International Carnahan Conference on Security Technology (ICCST), IEEE, 2019.

[59] Lindauer, B., et al. "Generating Test Data for Insider Threat Detectors." *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, vol. 5.2, pp. 80–94, 2014.

[60] Glasser, J., and B. Lindauer. "Bridging the Gap: A Pragmatic Approach to Generating Insider Threat Data." The 2013 IEEE Security and Privacy Workshops, 2013.

[61] Koroniotis, N., et al. "Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-Iot Dataset." *Future Generation Computer Systems*, vol. 100, pp. 779–796, 2019.

[62] Jakobi, N., P. Husbands, and I. Harvey. "Noise and the Reality Gap: The Use of Simulation in Evolutionary Robotics." Advances in Artificial Life: Third European Conference on Artificial Life Granada, Spain, 4–6 June 1995.

[63] Koos, S., J.-B. Mouret, and S. Doncieux. "The Transferability Approach: Crossing the Reality Gap in Evolutionary Robotics." *IEEE Transactions on Evolutionary Computation*, vol. 17.1, pp. 122–145, 2012.

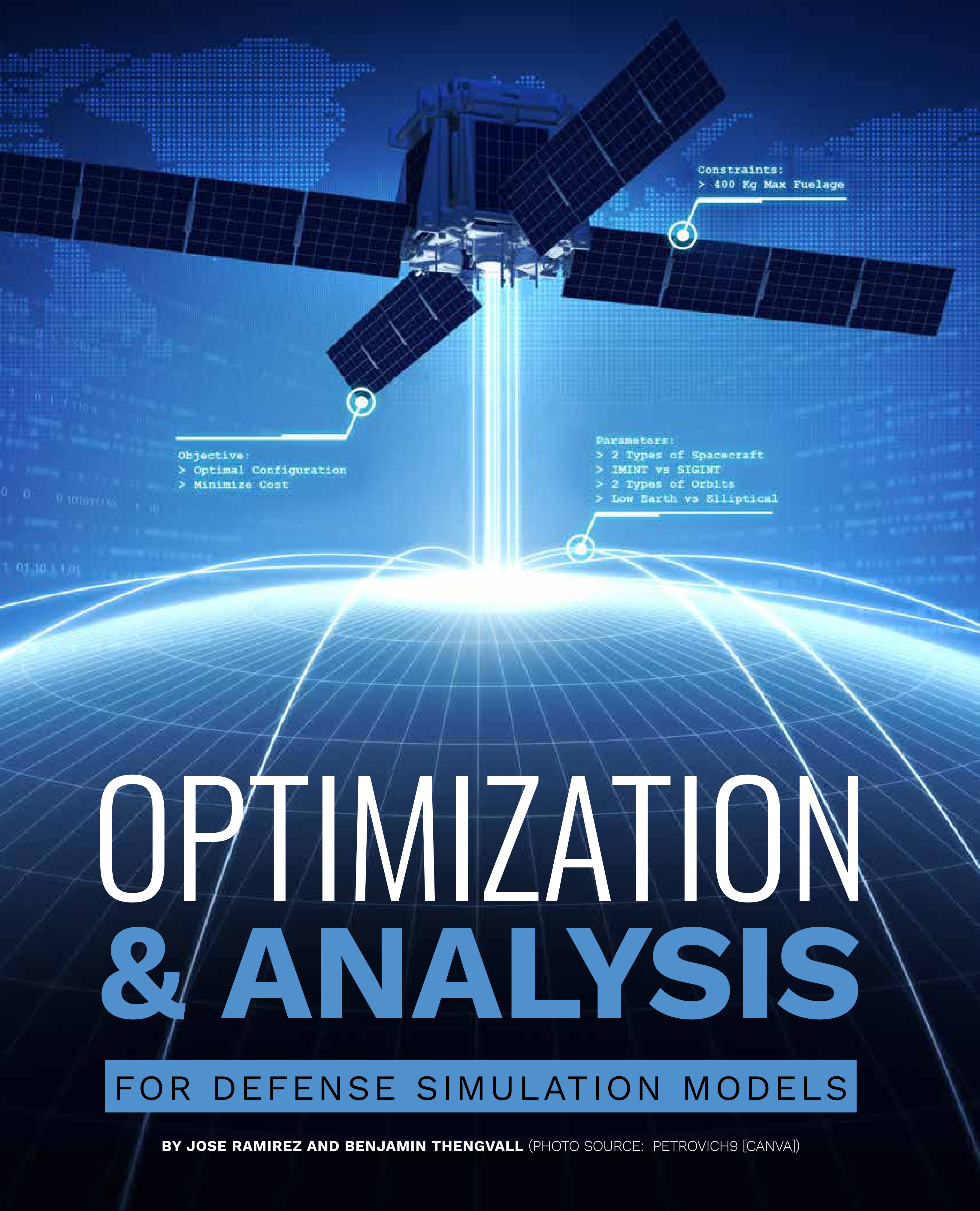
[64] Tene, O., and J. Polonetsky. "Privacy in the Age of Big Data: A Time for Big Decisions." *Stan. L. Rev. Online*, vol. 64, p. 63, 2011.

BIOGRAPHIES

EMILY NACK is an information technology specialist and research lab manager at the U.S. Army Cyber Institute within the United States Military Academy (USMA) at West Point. Her research focuses on cyber modeling and simulation (M&S), augmented reality for tactical and operational data visualization, and artificial intelligence integration in M&S environments. Ms. Nack holds a B.S. degree in game design and development from the Rochester Institute of Technology and an A.S. degree in computer information systems from Hudson Valley Community College.

NATHANIEL D. BASTIAN is a Lieutenant Colonel in the U.S. Army, where he is an academy professor and cyber warfare officer at Army Cyber Institute within the USMA at West Point and a program manager in the Information Innovation Office at the Defense Advanced Research Projects Agency. His research efforts aim to develop innovative, assured, intelligent, human-aware, data-centric, and decision-driven capabilities for multidomain operations. Dr. Bastian holds a B.S. in engineering management (electrical engineering) from USMA, an M.S. in econometrics and operations research from Maastricht University, and an M.Eng. in industrial engineering and Ph.D. in industrial engineering and operations research from Pennsylvania State University.

Photo Source: 123rf.com



OPTIMIZATION & ANALYSIS

FOR DEFENSE SIMULATION MODELS

BY JOSE RAMIREZ AND BENJAMIN THENGVALL (PHOTO SOURCE: PETROVICH9 [CANVA])

SUMMARY

When performing defense system analysis with simulation models, a great deal of time and effort is expended creating representations of real-world scenarios in U.S. Department of Defense (DoD) simulation tools. However, once these models have been created and validated, analysts rarely retrieve all the knowledge and insights that the models may yield and are limited to simple explorations because they do not have the time and training to perform more complex analyses manually. Additionally, they do not have software integrated with their simulation tools to automate these analyses and retrieve all the knowledge and insights available from their models.

Simple, manual explorations are inefficient in their use of computing resources and often ineffective in providing the best answers to analyst questions. To derive the greatest benefit from a simulation model, analysts should apply optimization and statistical analysis techniques. Combining these techniques and using available simulation optimization and analysis tools can provide answers to these essential questions and key insights for decision-makers. More importantly, the organizational return on investment from simulation studies increases, which builds stakeholder confidence. Tools like these can also be used for model verification and validation.

INTRODUCTION

Simulation and optimization are two powerful methodologies widely used across the DoD [1–4]. Simulations are used to understand complex system behavior at multiple levels of fidelity. For example, simulation can be used to perform detailed engineering, design, and testing of individual weapons system components; examine the interaction of components in a single advanced weapon system such as a fighter aircraft or nuclear submarine; and analyze tactical engagements between weapons systems.

Optimization provides a powerful way to determine the best option among many options. For example, military analysts use optimization to maximize the amount of fuel and munitions delivered to an area of operations using the least number of ships and cargo aircraft; determine the best allocation of dollars to minimize the risk of failure in a future war; assign military personnel to bases in a way that maximizes personal preferences and professional development; and allocate blue force weapons to red force targets to maximize the probability of damage while minimizing collateral damage. When resources are limited and mission effectiveness is paramount, optimization offers military decision-makers keen insights and enables them to make the best choices.

The core challenge in simulation optimization is finding the "best options" within environments too complex or uncertain for traditional optimization techniques. Due to their ability to handle these unpredictable factors, simulations are often the only way to model such problems. However, this creates a dilemma—simulation models become necessary because traditional optimization methods fail under these conditions. The very complexity built into the simulation model makes finding optimal solutions a daunting task. Until recently, no search process was sophisticated enough to bridge this gap between the power of simulation and the structured goal-finding nature of optimization. In short, no type of search process exists that can effectively integrate simulation and optimization. The same shortcoming is also encountered in settings outside of simulation where complex (realistic) models cannot be analyzed using traditional "closed form" optimization tools like mathematical programming.

Recent developments are changing this picture. Advances in metaheuristics—the domain of optimization that incorporates artificial intelligence and analogs to physical, biological, or evolutionary processes—have led to creating a new approach that successfully integrates simulation and optimization. As a result, analysts can get the best benefits from their simulation models.

Organizations may fail to take full advantage of their simulation models. Even though large amounts of time and money are invested in creating a simulation tool and populating it with validated data, a large part of the valuable knowledge that the model may yield is generally overlooked. Simulation analysts who can access such knowledge are exceedingly valuable to their organization and become highly sought-after resources. Combining optimization and statistical analysis techniques with a simulation model is key to unlocking this knowledge. Optimization techniques can be used to execute a simulation model many times, varying the input parameter values, to determine the best input values to achieve desired system outputs. The results of these simulation runs can then be explored with statistical techniques to better understand the system modeled by the simulation. Essential optimization and analysis questions that can be answered for simulation models by combining these techniques include the following:

> **“**
>
> ***Advances in metaheuristics have led to creating a new approach that successfully integrates simulation and optimization.***

- Optimization
  - What combinations of input parameters lead to the best and worst performance of the system?
  - What are the best tradeoffs between multiple competing objectives?
- Analysis
  - Which input parameters have the greatest influence on the system being modeled, and which have the least?
  - Are there good or bad regions of the input parameter space that can be defined by a subset of input parameters with restricted ranges?
  - Are some areas of the parameter trade space more robust to parameter variation than others?

To derive the greatest benefit from a simulation model, an analyst should apply optimization and statistical analysis techniques. Combining these techniques can provide answers to these essential questions and key insights for decision-makers. More importantly, they increase the organizational return on investment from their analysts and simulation models, such as providing a range of force structure capacity (size) options.

In this article, some of the most relevant approaches developed for optimizing simulated systems are summarized. The metaheuristic black-box approach that leads practical applications and relevant details of how this approach has been implemented and used in commercial software is provided next. As a concrete example, some of the mathematics and logic behind a generic simulation optimization software engine are described. Lastly, some use cases that analysts might encounter are presented, and how using a simulation optimization and analysis tool integrated with their simulation model can lighten their workload and lead to better study results is discussed.

OPTIMIZATION AND STATISTICAL ANALYSIS IN COMMERCIAL SIMULATION PACKAGES

Over the past two decades, optimization tools in commercial simulation packages have become widespread and relatively easy to use, even if not all practitioners exploit them. Commercial simulation packages also have analysis tools that explore the variability uncovered through simulation replications (or Monte Carlo runs) for a single set of input parameters. However, the analysis of all simulation runs resulting

> ***Over the past two decades, optimization tools in commercial simulation packages have become widespread and relatively easy to use.***

from an optimization run is less commonly available, at least in an automated, easy-to-digest way.

The underlying statistical techniques discussed in this article are not new. However, in many tools today, to perform variable sensitivity and good and bad region analysis across simulation runs executed with different combinations of input parameter changes, analysts must use multiple tools or perform the simulations and then piece together the results of various statistical techniques. Therefore, these types of valuable simulation analyses are done infrequently and often performed only by technical consultants and advanced analysts. To perform them, users of discrete event simulation packages export their simulation results and then use specialized statistical tools like JMP or SPSS or write code in languages like R or Python for analysis. Users of spreadsheet-based Monte Carlo simulations have more statistical analysis tools at their disposal, but even for these analysts, gaining insights across all simulation runs is not an automated process.

The critical goals of identifying good and bad regions of a parameter trade space and discovering robust solutions are sometimes pursued by more advanced analysts through generating a response surface approximation by coupling design of experiments with simulation. This approximate response surface is then explored through various stochastic optimization techniques [5]. Such an approach generally relies on moving from tool to tool for the different steps in the process—generating the design of experiments, executing the simulations, and performing the stochastic optimization. This type of process has the conspicuous shortcoming of frequently oversimplifying complex response surfaces, which can entail a costly loss of valuable insights.

CLASSICAL APPROACHES FOR SIMULATION OPTIMIZATION

Fu [6] identifies the following four main approaches for optimizing simulations:

1. Stochastic approximation (gradient-based approaches)
2. Sequential response surface methodology
3. Random search
4. Sample path optimization (also known as stochastic counterpart)

Stochastic approximation algorithms attempt to mimic the gradient search method used in deterministic optimization. The procedures based on this methodology must estimate the gradient of the objective function to determine a search direction. Stochastic approximation targets continuous variable problems because of its close relationship with the steepest descent gradient search. However, this methodology has been applied to discrete problems [7].

Sequential response surface methodology is based on the principle of building local metamodels. The "local response surface" is used to determine a search strategy (e.g., moving in the estimated gradient direction), and the process is repeated. In other words, the metamodels do not attempt to characterize the response surface for the entire solution space but rather concentrate on the search's local area.

A random search method moves through the solution space by randomly selecting a point from the current point's neighborhood. This implies that a neighborhood must be defined as part of developing a random search algorithm. Random search has been applied mainly to discrete problems, and its appeal is based on existing theoretical convergence proofs. Unfortunately, these theoretical convergence results mean little in practice where it is more important to find high-quality solutions within a reasonable length of time than to guarantee optimum convergence in an infinite number of steps.

Sample path optimization is a methodology that exploits the knowledge and experience developed for deterministic continuous optimization problems. The idea is to optimize a deterministic function that is based on n random variables, where n is the size of the sample path. In the simulation context, the method of common random numbers is used to provide the same sample path to calculate the response over different values of the input factors. Sample path optimization owes its name to the fact that the estimated optimal solution that it finds is based on a deterministic function built with one sample path obtained with a simulation model. Generally, n needs to be large for the approximating optimization problem to be close to the original optimization problem [8].

Leading commercial simulation software employs metaheuristics as the methodology of choice to provide optimization capabilities to their analysts. This approach to simulation optimization is explored in the next section.

SIMULATION OPTIMIZATION APPROACH WITH METAHEURISTICS

Metaheuristics provide a way of considerably improving the performance of simple heuristic procedures. The search strategies proposed by metaheuristic methodologies result in iterative procedures that can explore solution spaces beyond the solution resulting from applying a simple heuristic. For example, genetic algorithms (GAs) and scatter search (SS) are population-based metaheuristics designed to operate on a set of solutions maintained from iteration to iteration. On the other hand, metaheuristics like simulated annealing and tabu search (TS) typically maintain only one solution by applying mechanisms to transform the current solution into a new one. Metaheuristics have been developed to solve complex optimization problems in many areas, with combinatorial optimization being one of the most fruitful. Very efficient procedures can be achieved by relying on context information, i.e., by taking advantage of specific information about the problem. The solution approach may be viewed as the result of adapting metaheuristic strategies to specific optimization problems. In these cases, there is no separation between the solution procedure and the model that represents the complex system.

> **_Metaheuristics provide a way of considerably improving the performance of simple heuristic procedures._**

Metaheuristics can be used to create solution procedures that are context independent, i.e., procedures capable of tackling several problem classes and not using specific information from the problem to customize the search. The original GA designs were based on this paradigm, where solutions to all problems were represented as a string of zeros and ones [9]. The advantage of this design is that the same solver can be used to solve a wide variety of problems because the *solver* uses strategies to manipulate the string of zeros and ones and a decoder is used to translate the string into a solution to the problem under consideration. The obvious disadvantage is that the solutions found by context-independent solvers might be inferior to those of specialized procedures when applying the same amount of computer effort (e.g., search time). Solvers that do not use context information are referred to as general-purpose or "black box" optimizers.

Figure 1 shows the black box approach to simulation optimization favored

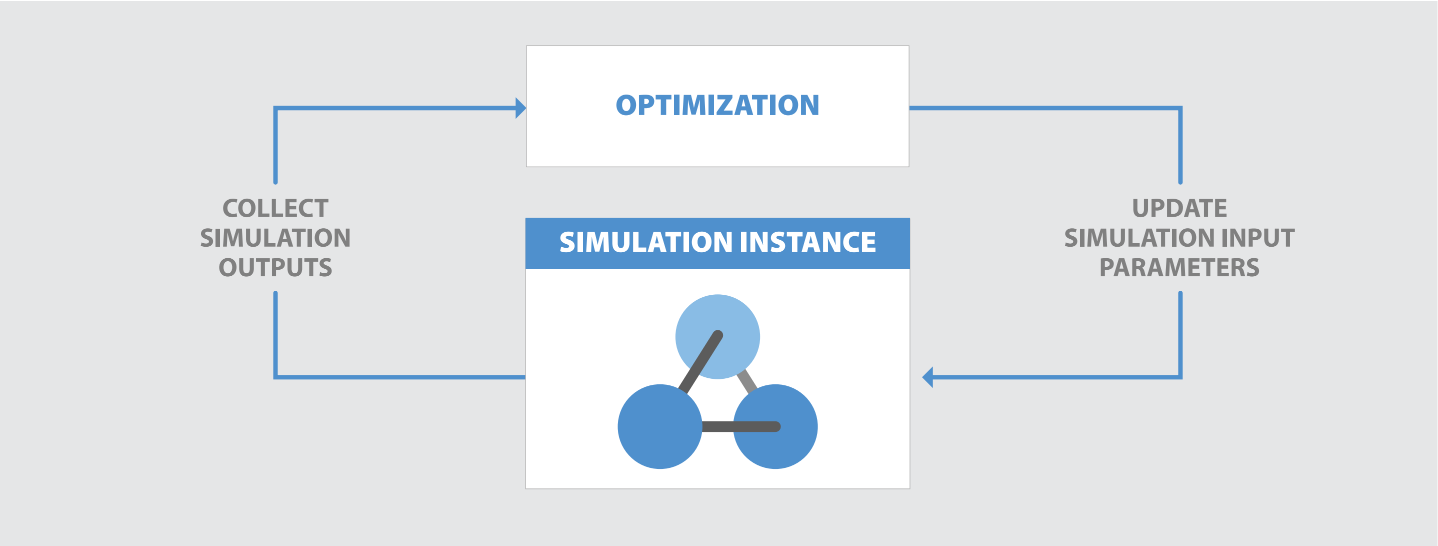

OPTIMIZATION
COLLECT SIMULATION OUTPUTS
SIMULATION INSTANCE
UPDATE SIMULATION INPUT PARAMETERS

Figure 1. Black Box Approach to Simulation Optimization *(Source: OptTek Systems)*.

by procedures based on metaheuristic methodology. In this approach, the metaheuristic optimizer (labeled as "optimization") chooses a set of values for the input parameters (i.e., factors or decision variables) and uses the responses generated by the simulation model or instance to make decisions for selecting the next trial solution.

One of the main design considerations when developing a general-purpose optimizer is which solution representation to employ. This representation is used to establish the communication between the optimizer and the simulation (which is the abstraction of the complex system). As previously mentioned, classical GAs used binary strings to represent solutions. This representation is not particularly convenient in some instances like when a natural solution representation is a sequence of numbers, as in the case of permutation problems. One of the most flexible representations is an n-dimensional vector, where each component may be a continuous or integer bounded variable. This representation can be used in a wide range of applications, which includes all those problems that can be formulated as mathematical programs.

SS

SS is a population-based metaheuristic for optimization. It has been applied to problems with continuous and discrete variables and with one or multiple objectives. The success of SS as an optimization technique is well documented in a constantly growing number of journal articles, book chapters [10–12], and a book [13]. SS consists of the following five phases:

1. Diversification Generation
2. Improvement
3. Reference Set Update
4. Subset Generation
5. Solution Combination

The Diversification Generation phase is used to generate a set of diverse solutions that are the basis for initializing the search. The Improvement phase transforms solutions to improve quality (typically measured by the objective function value) or feasibility (typically measured by some degree of constraint violation). The Reference Set Update phase refers to the process of building and maintaining a set of solutions that are combined in the main iterative loop of any SS implementation. The Subset Generation phase produces subsets of reference solutions which become the input to the combination method. The Solution Combination phase uses the output from the subset generation method to create new trial solutions. New trial solutions are the results of combining two or more reference solutions.

Extensions of the basic SS framework can be created to take advantage of additional metaheuristic search strategies, such as the memory-based structures of TSs. There are significant differences between classical GAs and SSs. While classical GAs rely heavily on randomization and limiting operations to create new solutions (e.g., one-point crossover on binary strings), SS employs strategic choices and memory, along with structured combinations of solutions, to create new solutions. SS explicitly encourages the use of additional heuristics to process selected reference points in search of improved solutions. This is especially advantageous in settings where heuristics that exploit the problem structure can either be developed or are already available.

OPTIMIZATION ENGINES

Many commercial and open-source simulation optimization engines exist. These engines often implement a composite of prediction and optimization technologies to tackle complex problems. They are particularly well-suited for scenarios where evaluating the objective function of the problem is computationally expensive. These engines utilize prediction models to help guide the search and estimate objective function values before solutions are evaluated. Commercial examples include OptQuest, OptDef, Simulink Design Optimization, and Hexaly. Open-source examples include the Python libraries ParMOO and RPFOpt.

Presented in this article are examples and use cases of a commercial solution that implements SS in the simulation optimization engine. This solution has been built under the following assumptions:

- A computationally expensive black box is used to evaluate the objective function of the optimization problem being solved.
- Prediction models within the engine have the dual purpose of assisting in establishing search directions and estimating the value of the objective function before solutions are processed by the objective function evaluator.

Prediction Technologies

Optimization engines often include multivariate linear regression modules to assess the linearity of unknown objective functions. If a reasonably accurate linear approximation can be obtained, this module may help filter out trial solutions unlikely to yield improvements before they are submitted for full evaluation, thus saving computational resources.

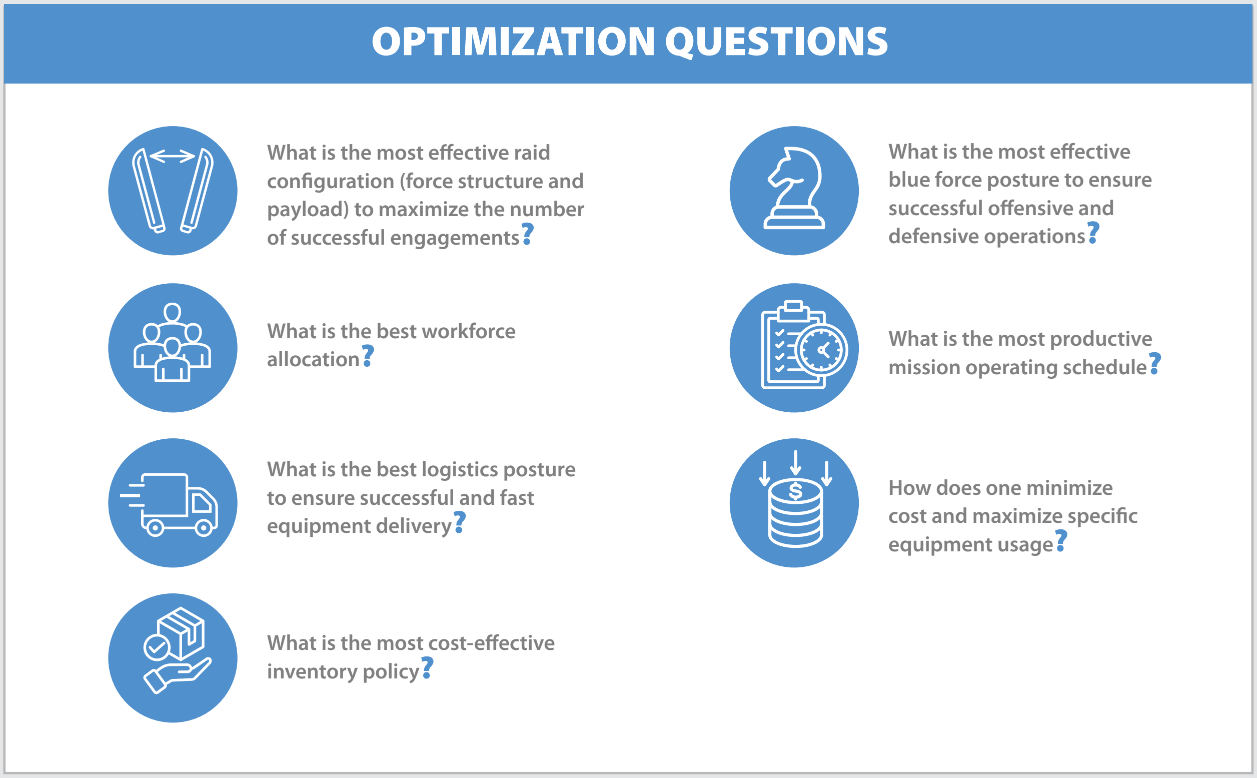
Neural networks are another prediction technology used in some optimization engines. These networks can be trained on already evaluated solutions to predict inferior trial solutions as well as suggest promising, high-quality solutions for subsequent evaluation.

Optimization Engine Capabilities and Practical Implications

Optimization engines have the potential to replace manual trial-and-error or basic parametric search methods, providing a more efficient way to identify promising decisions within a simulated or modeled domain. This is particularly valuable in defense simulations where analysts often lack tools to guide the selection of alternatives that yield optimal decisions. Figure 2 lists some optimization questions that are relevant to defense analysts.

Efficiently answering these questions often requires evaluating a massive number of scenarios through simulation or modeling tools. Optimization engines can automate

> ***Optimization engines have the potential to replace manual trial-and-error or basic parametric search methods, providing a more efficient way to identify promising decisions within a simulated or modeled domain.***



OPTIMIZATION QUESTIONS

What is the most effective raid configuration (force structure and payload) to maximize the number of successful engagements?

What is the best workforce allocation?

What is the best logistics posture to ensure successful and fast equipment delivery?

What is the most cost-effective inventory policy?

What is the most effective blue force posture to ensure successful offensive and defensive operations?

What is the most productive mission operating schedule?

How does one minimize cost and maximize specific equipment usage?

Figure 2. Optimization Questions for Defense Analysts *(Source: gravisio, Uniconlabs, Pure Template, Pexelpy, Iconbunny, and oksanavectorart Canva).*

the search for the best solutions. They enable decision-makers to define constraints, such as the following:

- Ranges of key parameters
- Budget limitations
- Asset capacities
- Acceptable minimum and maximum output values
- Limits on resources used
- Links between components or subsystems

The optimization engine strategically explores options within these constraints and then determines the strategic options investigated under its guidance, which it successively passes to the simulation or technical model for evaluation. The resulting search isolates scenarios that yield the highest quality outcomes for provided objectives, according to the criteria selected by the decision-maker.

USE CASE EXAMPLES

When coupled with appropriate defense simulation models, a simulation optimization and analysis tool enables optimization in various real-world scenarios, such as developing and refining concepts of operation, optimizing air defense configurations, maximizing satellite coverage, performing cybersecurity vulnerability assessments, and launching and deploying hypersonic weapons.

- **Optimal Blue Response.** Figure 3 shows a notional Advanced Framework for Simulation, Integration, and Modeling (AFSIM) scenario. Incoming blue forces attempt to hit all red targets. In this case, an analyst's objective would be to maximize the number of hits

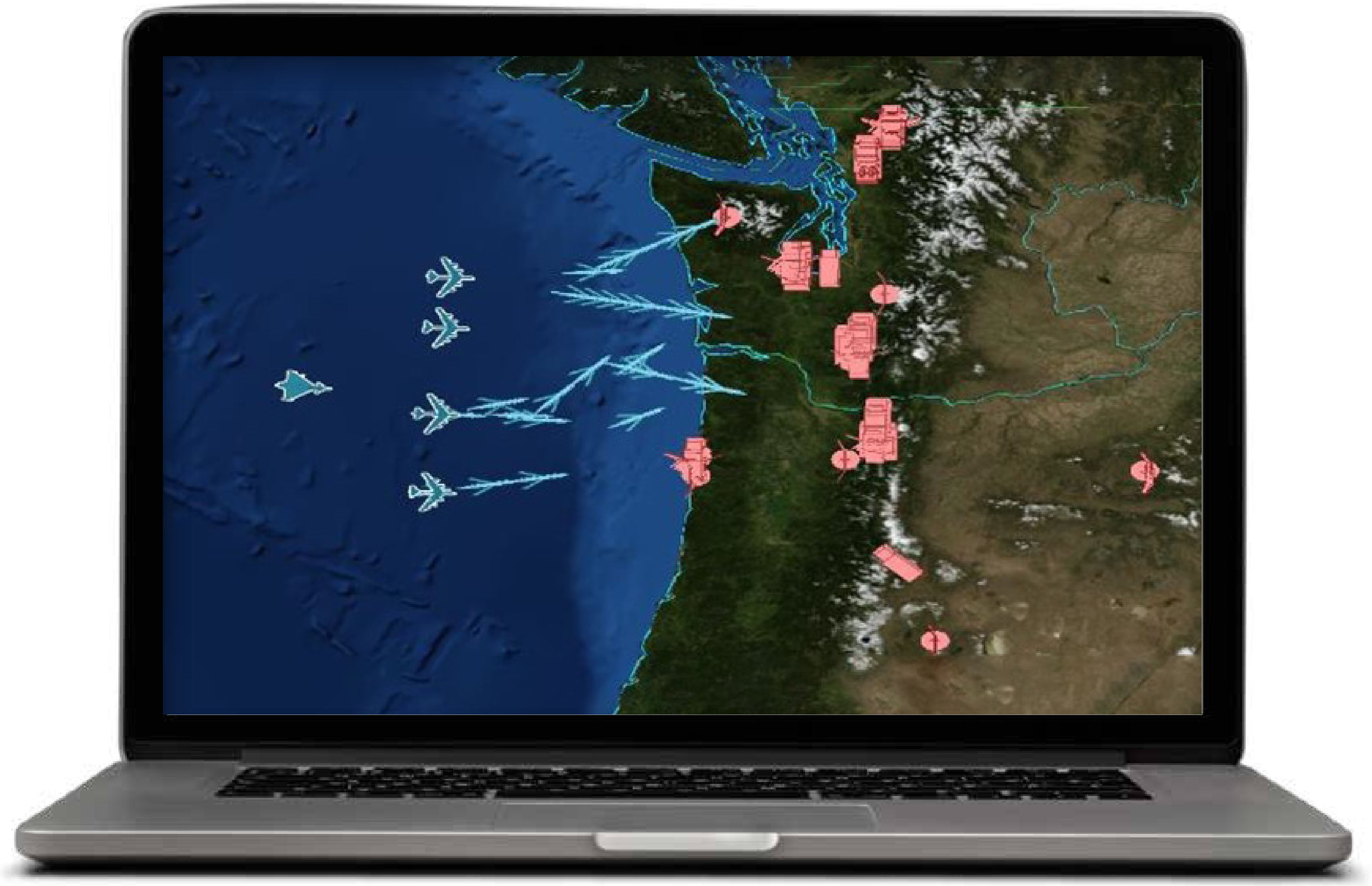


Figure 3. Notional AFSIM Scenario *(Source: OptTek Systems, Billion Photos [Canva]).*

while minimizing the number of blue aircraft used. Using the least number of blue forces has the added benefit of cost savings while still achieving the mission objective. A typical optimization setup would include varying parameters such as weapon type (categorical variable type), number of weapons (integer variable type), and the amount of time each aircraft has on a target (integer variable type). Running the simulation optimization software utilizes the metaheuristic methodologies described earlier to explore the space and find the optimal response (i.e., reach the analyst's objective).

- **Maximal Satellite Coverage.** An analyst may need to optimize satellite target coverage (i.e., swath). In this case, the objective could be finding the optimal satellite configuration to get the best coverage for high-priority targets/areas. An optimization setup could include varying spacecraft orbital parameters and system configuration (e.g., varying orbits, number, and type of spacecraft) to reach the objective. The analyst can also specify multiple objectives that would balance the best satellite configuration with cost (perhaps an additional variable would be fuel/energy amount). Utilizing multiple objectives allows the analyst to make data-driven decisions that best meet mission needs.

- **Cybersecurity Optimization and Analysis.** An analyst may want to test the limits of the information technology system by conducting vulnerability assessments. Unlike the Figure 3 scenario, which has blue forces on the offensive, the cybersecurity realm focuses on a defensive posture. Variable parameters may include number and type of servers that are part of the system architecture. It may also include known speed of response to a detected threat. If there is a cyberattack on the system, the analyst can optimize the solution to minimize loss of function and duration of effect against it. With simulation optimization, the analyst can explore scenarios in the cyber kill chain that are the most detrimental and identify key components that must be protected at all costs.

CONCLUSIONS

The fundamental principles of simulation optimization, from established research approaches to the metaheuristic strategies common in commercial applications, have been explored. Key implementation considerations, such as solution representation, metamodel utilization, and constraint formulation, were highlighted.

The synergy between simulation and optimization unlocks a level of solution quality far beyond manual "what-if" analysis, especially when the number of possible scenarios is vast. An overview of a commercial solution's optimization engine was provided, and the potential of simulation optimization tools to tackle complex defense-related problems when paired with simulation models was demonstrated.

Simulation optimization remains a vibrant field of research and development. Its versatility across diverse applications and the significant benefits it offers ensure continued advancements. Simulation optimization tools provide analysts with powerful resources to analyze complex systems and make data-driven decisions that optimize project and mission objectives. There is still much to learn and discover about how to optimize simulated systems from the theoretical and practical points of view. The rich variety of practical applications and the dramatic gains already achieved by simulation optimization ensure that this area will provide an intensive focus for study and a growing source of practical advances in the future. Simulation optimization software and tools can provide great support to analysts as they explore their data and achieve project/mission objectives. ■

> **_Simulation optimization tools provide analysts with powerful resources to analyze complex systems and make data-driven decisions that optimize project and mission objectives._**

REFERENCES

[1] Boginski, V., E. L. Pasiliao, and S. Shen. "Special Issue on Optimization in Military Applications." *Optimization Letters*, vol. 9, no. 8, pp. 1475–1476, 2015.

[2] Dirik, N., S. N. Hall, and J. T. Moore. "Maximizing Strike Aircraft Planning Efficiency for a Given Class of Ground Targets." *Optimization Letters*, vol. 9, no. 8, pp. 1729–1748, 2015.

[3] Kannon, T. E., S. G. Nurre, B. J. Lunday, and R. R. Hill. "The Aircraft Routing Problem With Refueling." *Optimization Letters*, vol. 9, no. 8, pp. 1609–1624, 2015.

[4] Hill, R. R., J. O. Miller, and G. A. McIntyre. "Simulation Analysis: Applications of Discrete Event Simulation Modeling to Military Problems." Winter Simulation Conference Proceedings, Arlington, VA, IEEE Computer Society, 2001.

[5] Samuelson, D. "When Close Is Better Than Optimal." *ORMS-Today*, vol. 37, no. 6, pp. 144–152, 2010.

[6] Fu, M. "Optimization for Simulation: Theory and Practice." *INFORMS Journal on Computing*, vol. 14, no. 3, pp. 192–215, 2002.

[7] Gerencser, L., S. D. Hill, and Z. Vago. "Optimization Over Discrete Sets via SPSA." *Proceedings of the 38th IEEE Conference on Decision and Control*, Phoenix, AZ, vol. 2, pp. 1791–1795, 1999.

[8] Andradottir, S. "A Review of Simulation Optimization Techniques." Proceedings of the 1998 Winter Simulation Conference, D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan (editors), pp. 151–158, 1998.

[9] Katoch, S., S. S. Chauhan, and V. Kumar. "A Review on Genetic Algorithm: Past, Present, and Future." *Multimedia Tools and Applications*, vol. 80, pp. 8091–8126, https://doi.org/10.1007/s11042-020-10139-6, 2021.

[10] Glover, F., M. Laguna, and R. Martí. *Scatter Search: Advances in Evolutionary Computation: Theory and Applications*, New York, NY: Springer-Verlag, 2003.

[11] Glover, F., M. Laguna, and R. Martí. *Scatter Search and Path Relinking: Advances and Applications*: Handbook of Metaheuristics, Boston, MA: Kluwer Academic Publishers, 2003.

[12] Glover, F., M. Laguna, and R. Martí. "New Ideas and Applications of Scatter Search and Path Relinking." *New Optimization Techniques in Engineering*, Berlin, Germany: Springer, 2004.

[13] Laguna, M., and R. Martí. *Scatter Search: Methodology and Implementations in C*. Boston, MA: Kluwer Academic Publishers, 2003.

BIOGRAPHIES

JOSE RAMIREZ is vice president of government services at OptTek Systems, Inc. He served as the Warfighting Analysis Division Chief (Colonel) in the Joint Staff J-8, where he led, managed, and provided analytical guidance to a 40-personnel team of analytical modelers, national defense strategists, and cybersecurity personnel. He has guided campaign-level, state-of-the-art, discrete-event modeling and simulation and data analytics supporting the Chairman of the Joint Chiefs of Staff and Combatant Commanders. He has strategized and conducted operational assessments of capabilities vs. peer adversaries for insights on future equipment modernization investment options and implemented advanced data analytics technology on the DoD's global munitions requirements process. Dr. Ramirez holds a B.S. in civil engineering from the University of Notre Dame, an M.S. in government information leadership from the National Defense University's College of Information and Cyberspace, an M.S.E. in operations research and industrial engineering from the University of Texas at Austin, and a Ph.D. in operations management (operations research) from the University of Colorado at Boulder.

BENJAMIN THENGVALL is chief operating officer at OptTek Systems, Inc. He is an expert in mathematical modeling, real-time optimization software and services, transportation and scheduling problems, agent-based and discrete-event simulation, and simulation optimization and analysis. He has spent his career providing innovative software solutions to complex real-world problems through mathematical modeling, simulation, and metaheuristic techniques in commercial and government spheres. Dr. Thengvall holds a B.S. in mathematics from the University of Nebraska-Lincoln and an M.S.E. and Ph.D. in operations research and industrial engineering from the University of Texas at Austin.

HAVE AN IDEA FOR
AN ARTICLE?
If you would like to publish with CSIAC or have an idea for an article, we would love to hear from you.
To learn more, visit www.csiac.mil/publish

Photo Source: Katerina Holmes (Canva)

BY MAZAHER KIANPOUR

(PHOTO SOURCE: GURT.SPACE, PIXABAY [CANVA])

ARTIFICIAL INTELLIGENCE IN CYBERSECURITY

HOW SUBSTITUTING AND SCALING IMPACT INVESTMENT RETURNS

SUMMARY

With the growing integration of artificial intelligence (AI) in cybersecurity, this article investigates the economic principles of substitution and scale's elasticity to evaluate their impact on the return on security investment. Recognizing the potential of AI technologies to substitute human labor and traditional cybersecurity mechanisms and the significance of cost ramifications of scaling AI solutions within cybersecurity frameworks, the study theoretically contributes to understanding the financial and operational dynamics of AI in cybersecurity. It provides valuable insights for cybersecurity practitioners in public and private sectors. Through this analysis, ways in which AI technologies can redefine economic outcomes in cybersecurity efforts are highlighted. Strategic recommendations are also offered for practitioners to optimize the economic efficiency and effectiveness of AI in cybersecurity, emphasizing a dynamic, nuanced approach to AI investment and deployment.

INTRODUCTION

> ***In the rapidly evolving landscape of cybersecurity, the integration of artificial intelligence (AI) technologies represents a paradigm shift.***

In the rapidly evolving landscape of cybersecurity, the integration of artificial intelligence (AI) technologies represents a paradigm shift, offering unprecedented opportunities for enhancing security measures against complex cyberthreats [1, 2]. This transition, driven by the increasing sophistication of cyberattacks and extremely large and diverse collections of structured and unstructured data generated in digital ecosystems, necessitates a reevaluation of traditional cybersecurity frameworks. As such, governments recognize the imperative to adapt and innovate, leveraging AI to strengthen and improve defense mechanisms [3] and ensure efficient utilization of resources in public security domains [4].

Building on this foundation, the economic principles of elasticity of substitution and elasticity of scale emerge as pivotal factors in this context, providing a lens through which the impact of AI on cybersecurity can be assessed for operational efficiency and investment return. The elasticity of substitution explores the extent to which AI technologies can replace traditional human labor and non-AI cybersecurity measures. The elasticity of scale examines the cost implications of scaling AI solutions within cybersecurity infrastructures. This scaling is especially important in public sector contexts where budget constraints are ongoing issues and maximizing resource efficiency is crucial [5].

Given the strategic importance of cybersecurity investments, understanding these elasticities' influence on the return on security investment (ROSI) is crucial for organizations, including government entities, navigating the digital transformation. This article aims to explore the interplay between the elasticity of substitution, the elasticity of scale, and ROSI in the context of AI in cybersecurity, framing an analysis that aids in strategic decision-making for investments in cybersecurity technologies. The core research question guiding this exploration is: How do the elasticities of substitution and scale influence ROSI? Addressing this question is significant for several reasons.

Firstly, it contributes to the literature at the intersection of economic theory and cybersecurity, offering an analytical framework for understanding the financial and operational dynamics of AI integration. Secondly, by examining the economic implications of deploying AI in cybersecurity, the study provides practical insights for businesses, policymakers, and cybersecurity professionals, facilitating informed decisions that balance cost, efficiency, and security outcomes.

Lastly, the investigation into ROSI underscores the financial viability of AI cybersecurity solutions, a critical concern for stakeholders in an era of tightening budgets and escalating cybersecurity risks.

This article is structured as follows. The Theoretical Foundations section presents the theoretical foundation of this article by exploring the concepts of return on security investment along with the elasticities of substitution and scale. The Analysis section is dedicated to presenting an analysis. Following this, the Implications and Suggested Strategies for Practitioners section highlights the implications and proposes strategies for the cybersecurity practitioners. Finally, the Conclusions section concludes the article.

THEORETICAL FOUNDATIONS

In digital ecosystems, AI does not just automate tasks typically reserved for low-skilled labor but is also involved in domains once thought to be exclusive to high-skilled labor through its innovative capabilities [6]. For instance, in cybersecurity, AI algorithms are not only replacing routine tasks like malware detection but are also stepping into roles requiring complex decision-making, such as identifying subtle patterns of sophisticated cyberattacks or automating the response to incidents in real-time. This leap signifies a shift from AI as a tool for automation to a comprehensive strategic asset capable of driving innovation in cybersecurity measures [7]. This transformative potential of AI in cybersecurity directly ties into the study's examination of the elasticity of substitution and scale, highlighting the economic impacts of integrating AI into security strategies.

The elasticity of substitution emerges as a critical factor when considering the replacement of human labor and non-AI cybersecurity measures with AI-driven tools. This elasticity measures the ease with which AI technologies can be substituted for traditional security methods, influenced by the technological advancement of AI, its compatibility with existing security infrastructures, regulatory compliance requirements and accountability measures, and the dynamic nature of cyberthreats. The relative costs of labor and AI technologies play a significant role in this dynamic, where advancements in AI capabilities and labor market fluctuations can shift the balance, potentially making AI solutions more economically attractive [8]. Such cost shifts can accelerate or hinder the adoption of AI in cybersecurity, reflecting on the broader implications for security effectiveness and organizational resilience.

> ***AI algorithms are not only replacing routine tasks like malware detection but are also stepping into roles requiring complex decision-making.***

Parallelly, the elasticity of scale addresses how the expansion of AI in cybersecurity affects cost structures, focusing on the implications of scaling AI solutions for the overall economics of cybersecurity efforts. By their automated and digital nature, AI technologies present unique opportunities for economies of scale, where the marginal cost of cybersecurity operations could decrease as AI solutions are deployed more extensively. However, this optimistic view is balanced by considering possible diseconomies of scale, such as the added complexity and overhead that might accompany large-scale AI deployments, potentially eroding the cost benefits of scalability.

A pertinent real-world example that illustrates the elasticity of scale in AI-driven cybersecurity is the implementation of AI technologies to enhance software supply chain (SSC) security within the Defense Industrial Base (DIB). This approach, as detailed in the report published by the Cybersecurity & Information Systems Information Analysis Center (CSIAC) [9], leverages AI to automate threat intelligence processing and expedite

> ***A pertinent real-world example that illustrates the elasticity of scale in AI-driven cybersecurity is the implementation of AI technologies to enhance software supply chain security within the Defense Industrial Base.***

cybersecurity risk management. As these AI systems scale across the defense industry's complex infrastructure, they demonstrate economies of scale by spreading the development and operational costs over a larger network of military and government installations. For instance, AI-driven systems can provide continuous scanning and real-time threat assessment across various platforms, significantly reducing the marginal cost of enhancing security for each additional system component.

However, the scalability and cost-effectiveness of these AI-driven solutions can be tempered by diseconomies of scale as they grow. For example, as AI solutions are scaled up to protect SSCs across the DIB, which comprises various branches of the military with distinct operational environments, the complexity of ensuring compliance with the National Institute of Standards and Technology controls adds significant overhead. According to the CSIAC report [9], integrating and managing AI-driven threat assessment tools across different branches involves substantial costs related to system customization to adhere to specific security standards, training personnel to handle sophisticated AI tools, and updating systems to keep up with the latest security protocols. These complexities can lead to slower response times to emerging threats and increase the operational costs, thereby potentially diluting the initial cost benefits associated with scaling AI solutions in such a regulated and diverse environment.

This study aims to investigate the impact of elasticity of substitution and scale on the ROSI, given that ROSI is profoundly affected by the cost-effectiveness and efficiency of cybersecurity measures. By integrating the effects of substitution and scale, understanding how strategic AI integration can significantly enhance ROSI is explored. This metric has been extensively studied in the literature of cybersecurity economics [10, 11]. ROSI provides a quantitative measure of the financial value derived from implementing cybersecurity countermeasures [12], serving as a pivotal tool for assessing the economic viability of investments in cybersecurity technologies, including AI. One method to quantitatively calculate ROSI is as follows [13]:

$$ROSI = \frac{Cost\ Avoided - Cost\ of\ Investment}{Cost\ of\ Investment} \times 100\%, \quad (1)$$

where:

- *Cost Avoided* (*CA*) includes potential losses from cybersecurity incidents that are prevented due to AI-enhanced security measures. It can also consider saved costs and efficiency gains, such as reduced downtime or faster threat detection and response times.

- *Cost of Investment* (*CI*) encompasses the total expenditure on AI technologies, including initial purchase, implementation, training, and ongoing maintenance costs.

Understanding the elasticity of substitution sheds light on the potential of AI to replace existing cybersecurity measures efficiently, potentially leading to a substantial reduction in CA due to risk mitigation and response capabilities. Concurrently, examining the elasticity of scale allows for an assessment of how the costs associated with AI-driven cybersecurity solutions evolve as these solutions are deployed at larger scales, affecting the overall ROSI calculation. Through this lens, the next section explores the complex interplay between these elasticates and ROSI, highlighting the nuanced ways in which AI technologies can redefine economic outcomes in cybersecurity efforts.

ANALYSIS

This section presents an analysis to elucidate the theoretical impact of elasticity of substitution and elasticity of scale on ROSI outcomes within the framework of AI-driven cybersecurity measures. Integrating these foundational economic theories to construct a comprehensive understanding of how the adaptability and scalability of AI technologies influence their economic viability and effectiveness in enhancing cybersecurity defenses is the goal. The elasticity of substitution is a measure of how easily one (economic) good or input can be substituted for another in response to changes in their relative prices or productivity [14]. The following widely applicable formula is adapted for calculating the elasticity of substitution [15] to better fit the unique context and specific requirements of the analysis:

$$\sigma_{A,X} = \frac{d(\ln\left(\theta(A).\frac{A}{X}\right))}{d(\ln\left(\frac{MP_A}{MP_X}\right))}. \tag{2}$$

In this equation, *A* and *X* represent the inputs of AI investment and traditional inputs (labor and non-AI technologies), respectively. $\theta(A)$ captures the efficiency or effectiveness of AI technology as a function of AI investment (*A*), reflecting how advancements in AI technology improve its contribution to cybersecurity effectiveness. Technological advancements in AI ($\theta(A)$) can alter the marginal productivity of AI (MP_A), potentially increasing its substitutability with traditional inputs (*X*). This enhanced substitutability, driven by AI's technological advancements, affects both the cost avoided and the cost of investment, thereby influencing ROSI.

The elasticity of scale examines how the total output, in this case, cybersecurity effectiveness (*Y*), changes in response to a proportional increase in all inputs (*A* and *X*) [16]. This analysis is crucial for understanding the scalability of AI investments in cybersecurity and their potential to yield returns to scale. Identifying the conditions under which AI investments lead to enhanced scalability of cybersecurity operations can inform strategic decisions about the pace and extent of AI integration. The elasticity of scale, modified to account for external factors (ϕ), examines how the scalability of cybersecurity effectiveness is influenced by these factors as follows:

$$\varepsilon = \frac{d(ln Y)}{d(ln\lambda)} \cdot \frac{d\Psi(\phi)}{d\phi}. \tag{3}$$

This equation measures how the output (*Y*) changes in response to a proportional change in all inputs, scaled by a factor of λ. Moreover, it indicates that the overall returns to scale in cybersecurity effectiveness can be affected by external factors, which can either amplify or diminish the effectiveness of scaling up inputs, including AI. $\Psi(\phi)$ is a function that modifies the overall effectiveness of the cybersecurity system based on external factors. In this study, it is assumed that $\Psi(\phi)$ increases with positive external developments (e.g., effective AI regulations) and decreases with negative developments (e.g., sophisticated cyberthreats).

These two refined metrics, encompassing technological advancements in AI ($\theta(A)$) and the role of external factors (ϕ), delineate four distinct scenarios that significantly influence the ROSI in the domain of AI-driven cybersecurity. Each scenario represents a unique combination of the elasticity of substitution and scale, providing a nuanced understanding of how AI's integration into cybersecurity strategies can be optimized for economic efficiency and effectiveness, as follows:

1. High Elasticity of Substitution ($\sigma_{A,X}$>1) With High Elasticity of Scale (ε>1)

This scenario signifies an ideal state where technological advancements in AI not only enhance its substitutability with traditional cybersecurity measures but also facilitate economies of scale as AI solutions are expanded. The dual presence of a high $\sigma_{A,X}$ due to significant $\theta(A)$ improvements, alongside a favorable ε influenced by positive external factors (ϕ), suggests an optimal environment for AI investments, yielding substantial improvements in ROSI. This combination reflects a scenario where

AI's deployment maximizes cost avoidance and minimizes investment costs, presenting a compelling case for aggressive AI integration in cybersecurity frameworks.

2. High Elasticity of Substitution ($\sigma_{A,X}$)>1) With Low Elasticity of Scale (ε<1)

In this scenario, while AI exhibits strong substitutability due to advancements ($\theta(A)$), scaling AI solutions encounters challenges, possibly due to negative externalities (ϕ) that diminish the returns to scale (ε<1). This juxtaposition leads to mixed outcomes for ROSI, where the benefits of substituting AI for traditional measures may be partially offset by the increased costs or diminished effectiveness associated with scaling. Strategic considerations must be employed to navigate this landscape, balancing the push for substitution with careful scaling strategies.

3. Low Elasticity of Substitution ($\sigma_{A,X}$)<1) With High Elasticity of Scale (ε>1)

Here, AI's technological advancements may not sufficiently enhance its substitutability, possibly due to limitations in AI's applicability or integration complexities. However, positive external factors support economies of scale, suggesting that while AI may not replace traditional measures as effectively, scaling up AI deployments is economically beneficial. This scenario requires a focused approach to leveraging the scalability of AI to improve ROSI, possibly by enhancing AI capabilities or finding niches where AI's integration delivers clear benefits.

4. Low Elasticity of Substitution ($\sigma_{A,X}$)<1) With Low Elasticity of Scale (ε<1)

Representing the most challenging scenario, this combination arises when AI's technological advancements fail to significantly increase its substitutability and external factors lead to diseconomies of scale. The convergence of these factors results in the lowest potential for ROSI improvement, indicating a need for a reevaluation of AI investment strategies. Organizations in this quadrant must critically assess their AI deployments, focusing on overcoming barriers to AI effectiveness and scalability to realize positive economic outcomes.

Through these scenarios, Table 1 represents a comprehensive matrix that captures the multifaceted impacts of AI's elasticity of substitution and scale on ROSI in cybersecurity.

Investigating the details of each scenario—identifying what factors make AI more scalable or substitutable to determine which scenario is more likely—is beyond the scope of this article and requires a specific, detailed analysis tailored to the distinct needs of each organization and their cybersecurity context. Nonetheless, based on this theoretical foundation, the next section draws practical implications of elasticity of substitution and scale on the economics of cybersecurity, guiding organizations in navigating the complexities of AI integration to optimize security investments and operational effectiveness.

Table 1. The Interplay Between Elasticities of Substitution and Scale on Return on Cybersecurity Investment

ELASTICITY OF SUBSTITUTION/SCALE	HIGH ELASTICITY OF SCALE (ε>1)	LOW ELASTICITY OF SCALE (ε>1)
High Elasticity of Substitution ($\sigma_{A,X}$>1)	High ROSI: Efficient substitution leads to significant CA, and increased returns to scale reduces CI per unit.	Mixed ROSI: While substitution is efficient, increasing CI due to diseconomies of scale could offset CA benefits.
Low Elasticity of Substitution ($\sigma_{A,X}$<1)	Mixed ROSI: Economies of scale lower CI per unit, but lower substitution efficiency reduces CA gains.	Low ROSI: Diseconomies of scale increase CI, and low substitution efficiency minimally impacts, resulting in minimal ROSI improvement.

IMPLICATIONS AND SUGGESTED STRATEGIES FOR PRACTITIONERS

Integrating the insights from the analysis of elasticity of substitution and elasticity of scale on ROSI, the practical implications can be synthesized into a cohesive strategy for cybersecurity practitioners. This strategy revolves around optimizing the economic efficiency and effectiveness of AI-driven cybersecurity measures by understanding and acting upon the interplay between these elasticities and ROSI. Figure 1 presents a consolidated action plan that reflects these insights.

These strategies together enable cybersecurity practitioners to construct a comprehensive approach that maximizes the economic efficiency and effectiveness of AI in cybersecurity. This approach acknowledges the complexity of the interplay between the elasticity of substitution and scale, guiding practitioners in making informed decisions that optimize ROSI in an ever-evolving cybersecurity landscape.

1 ASSESSMENT AND ALIGNMENT

Begin with a comprehensive assessment of current cybersecurity measures and AI technologies. Align AI investments with strategic security objectives, considering both the elasticity of substitution ($\sigma_{A,X}$) and elasticity of scale (ε) to identify areas where AI can either replace traditional methods (high $\sigma_{A,X}$) or augment them efficiently at scale (high ε).

2 DYNAMIC INVESTMENT SCALING

Adjust the scale of AI investments based on the elasticity of substitution and scale. In scenarios where both elasticities are high, scale investments aggressively to capitalize on both substitution and scale efficiencies. When elasticity of substitution is high but elasticity of scale is low, focus investments on high-impact, targeted AI applications that can replace or enhance specific cybersecurity tasks without the need for extensive scaling.

3 COST-EFFICIENCY OPTIMIZATION

Regularly evaluate the cost-efficiency of AI deployments in cybersecurity, considering both direct costs (e.g., licensing, development, and operation) and indirect costs (e.g., training and integration challenges). Optimize AI applications where the ratio of elasticity of substitution to the cost of implementation and scaling yields the highest ROSI.

4 MITIGATION OF NEGATIVE EXTERNALITIES

Identify and address the externalities (ϕ) that negatively impact the elasticities of AI applications in cybersecurity. This includes technological limitations, regulatory constraints, interoperability issues with existing systems, and scalability challenges. Develop and implement strategies to mitigate these barriers, such as investing in technology upgrades, advocating for regulatory changes, or enhancing AI integration capabilities.

5 HYBRID AI-HUMAN CYBERSECURITY MODELS

Recognize scenarios where low elasticity of substitution suggests AI cannot fully replace human tasks. In such cases, invest in hybrid models that leverage AI to augment human capabilities, focusing on scalability (high ε) to enhance overall cybersecurity posture. This approach can include AI-assisted threat detection and response, where AI algorithms identify potential threats and human experts make final determinations.

6 INCREMENTAL INNOVATION AND CONTINUOUS LEARNING

Embrace a culture of incremental innovation and continuous learning within a cybersecurity team. This involves staying abreast of the latest advancements in AI and cybersecurity, conducting pilot projects to evaluate new approaches, and learning from both successes and failures to refine an AI strategy over time.

7 COLLABORATION AND KNOWLEDGE SHARING

Engage in industry collaborations and knowledge-sharing initiatives to learn from the experiences of others and add insights. This can help in identifying best practices for leveraging AI in cybersecurity, understanding how other organizations have navigated the challenges of elasticity, and finding innovative solutions to common problems.

Figure 1. Consolidated Action Plan for Cybersecurity Practitioners *(Source: Visual Generation, 0721-Team, and Vir Leguizamón [Canva]).*

CONCLUSIONS

This article has provided a thorough analysis of the intricate dynamics between the elasticity of substitution and scale of AI technologies and their consequential impact on ROSI within the realm of cybersecurity.

> ***Integrating the insights from the analysis of elasticity of substitution and elasticity of scale on ROSI, the practical implications can be synthesized into a cohesive strategy for cybersecurity practitioners.***

Exploration revealed that the strategic integration of AI in cybersecurity is not merely a technological upgrade but a complex economic decision that hinges on understanding and leveraging the elastic properties of AI. The ideal scenario—characterized by high elasticity of substitution and scale—underscores the potential for AI to deliver substantial improvements in ROSI through cost-effective substitution and scalable deployment. However, the mixed and challenging scenarios present a call to action for organizations in public and private sectors to navigate the intricacies of AI deployment with agility and foresight, addressing barriers to scalability and enhancing substitutability where necessary. The interrelated nature of substitution and scale elasticities demands a dynamic, nuanced approach to AI investment and deployment in cybersecurity. Organizations must adopt a continuous evaluation mindset, recalibrating strategies in response to technological evolutions and shifting threat landscapes to harness AI's full economic and security potential. Moreover, this article calls for empirical evidence and case studies that illustrate the real-world applications and implications of these elasticities in cybersecurity strategies. ■

REFERENCES

[1] Kaur, R., D. Gabrijelčič, and T. Klobučar. "Artificial Intelligence for Cybersecurity: Literature Review and Future Research Directions." *Information Fusion*, 2023.

[2] Truong, T. C., I. Zelinka, J. Plucar, M. Čandík, and V. Šulc. "Artificial Intelligence and Cybersecurity: Past, Presence, and Future." *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, 2020.

[3] Mariarosaria, T., T. McCutcheon, and L. Floridi. "Trusting Artificial Intelligence in Cybersecurity Is a double-Edged Sword." *Nature Machine Intelligence*, pp. 557–560, 2019.

[4] Mikalef, P., L. Kristina, S. Cindy, Y. Maija, F. Siw Olsen, T. Hans Yngvar, G. Manjul, and N. Bjoern. "Enabling AI Capabilities in Government Agencies: A Study of Determinants for European Municipalities." *Government Information Quarterly*, 2022.

[5] Ponemon Institute. "State of Cybersecurity in Local, State & Federal Government," 2015.

[6] Aghion, P., B. F. Jones, and C. I. Jones. *Artificial Intelligence and Economic Growth*. Cambridge: National Bureau of Economic Research, 2017.

[7] Lu, Y., and Y. Zhou. "A Review on the Economics of Artificial Intelligence." *Journal of Economic Surveys*, 2021.

[8] Hakami, N. "Navigating the Microeconomic Landscape of Artificial Intelligence: A Scoping Review." *Migration Letters*, 2023.

[9] Rahman, A. "Applications of Artificial Intelligence (AI) for Protecting Software Supply Chains (SSCs) in the Defense Industrial Base (DIB)." Cybersecurity & Information Systems Information Analysis Center, 2024.

[10] Gordon, L. A., and M. P. Loeb. "Return on Information Security Investments: Myths vs. Realities." *Strategic Finance*, 2002.

[11] Sonnenreich, W., J. Albanese, and B. Stout. "Return on Security Investment (ROSI) - A Practical Quantitative Model." *Journal of Research and Practice in IT*, 2006.

[12] Kianpour, M., S. J. Kowalski, and H. Øverby. "Systematically Understanding Cybersecurity Economics: A Survey." *Sustainability*, 2021.

[13] Bistarelli, S., F. Fioravanti, P. Peretti, and F. Santini. "Evaluation of complex Security Scenarios Using Defense Trees and Economic Indexes." *Journal of Experimental & Theoretical Artificial Intelligence*, 2012.

[14] M. Knoblach, M., and F. Stöckl. "What Determines the Elasticity of Substitution Between Capital and Labor? A Literature Review." *Journal of Economic Surveys*, 2020.

[15] Robinson, J. *The Economics of Imperfect Competition*. London: Springer, 1933.

[16] Perloff, J. *Microeconomics*. Pearson Education, 2009.

BIOGRAPHY

MAZAHER KIANPOUR is a researcher with a deep interest in how technology and economics intersect, especially in cybersecurity and its economic and policy challenges. He is currently involved in research at RISE Research Institutes of Sweden and serves as a postdoctoral researcher at the Norwegian University of Science and Technology (NTNU), concentrating on the regulatory risks connected to cybersecurity. Dr. Kianpour holds a Ph.D. in information security from NTNU.

CSIAC Cybersecurity & Information Systems
Information Analysis Center

TECHNICAL INQUIRY SERVICES

FOUR FREE HOURS

Research within our four focus areas available to academia, industry, and other government agencies. Log in to csiac.mil to submit your inquiry today.

TECHNICAL AREAS

Cybersecurity

Knowledge Management & Information Sharing

Modeling & Simulation

Software Data & Analysis

Photo Source:
U.S. Air Force and 123.com

CS IAC JOURNAL

The Cybersecurity & Information Systems Information Analysis Center (CSIAC) is a component of the U.S. Department of Defense's (DoD's) Information Analysis Center (IAC) enterprise, serving the defense enterprise of DoD and federal government users and their supporting academia and industry partners.

WWW.CSIAC.MIL
CONNECT WITH US ON SOCIAL MEDIA.