JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2021, 1–13
doi: 10.1093/cybsec/tyab005
Review Article

Review Article

Simulation for cybersecurity: state of the art and future directions

Hamdi Kavak ,[1,]* Jose J. Padilla,[2] Daniele Vernon-Bido,[3] Saikou Y. Diallo,[2] Ross Gore,[2] and Sachin Shetty[2]

[1]Department of Computational and Data Sciences, George Mason University, 4400 University Drive, Fairfax, MS, 6A12, VA, USA; [2]Virginia Modeling, Analysis and Simulation Center, Old Dominion University, 1030 University Blvd., Suffolk, VA, 23435, USA and [3]Computational Modeling and Simulation Engineering, Old Dominion University, 1300 Engineering & Computational Sciences Building, Norfolk, VA, 23529, USA

*Correspondence address. Department of Computational and Data Sciences, George Mason University, 4400 University Drive, MS 6A12, Fairfax, VA 22030-4444, USA. Tel: +1–703–993–6376; E-mail: hkavak@gmu.edu

Received 1 October 2019; revised 29 April 2020; accepted 22 January 2021

Abstract

In this article, we provide an introduction to simulation for cybersecurity and focus on three themes: (1) an overview of the cybersecurity domain; (2) a summary of notable simulation research efforts for cybersecurity; and (3) a proposed way forward on how simulations could broaden cybersecurity efforts. The overview of cybersecurity provides readers with a foundational perspective of cybersecurity in the light of targets, threats, and preventive measures. The simulation research section details the current role that simulation plays in cybersecurity, which mainly falls on representative environment building; test, evaluate, and explore; training and exercises; risk analysis and assessment; and humans in cybersecurity research. The proposed way forward section posits that the advancement of collecting and accessing sociotechnological data to inform models, the creation of new theoretical constructs, and the integration and improvement of behavioral models are needed to advance cybersecurity efforts.

Key words: cybersecurity simulation, modeling and simulation, human representation in cybersecurity, cyber-physical system, cybersecurity modeling

Introduction

Reliance on information technology (IT) has grown significantly since the bloom of the Internet [1]. People and organizations use technology for mission-critical tasks such as banking, personnel management, or collaborative work. While IT makes accomplishing such tasks more convenient, it brings about serious security challenges that need to be addressed by all parties ranging from individuals to governments [2].

For individuals, one of the significant challenges is the release of personal information as a result of cyberattacks [3, 4]. Stolen identity information is mostly used for fraudulent transactions such as loan applications and tax returns. In 2010, there were an estimated 8.1 M victims of identity thefts in the USA alone [5]—by 2018, that number had risen to 14.4 M [6]. For corporations, financial losses due to cyberattacks are immense [7]. Lewis [8] estimates the annual costs of cybercrime to the global economy are getting close to US$500 billion, and this does not include the losses due to a damaged reputation. Furthermore, national security is also impacted by cyberattacks, targeting mission-critical private sector contractors and critical infrastructures, affect the stability of a country. In the USA, officials acknowledge that critical infrastructures have been under deliberate attacks, and repairing damages has been costly for the country [9, 10]. Furthermore, the USA is only one of the many victims of such attacks. In Estonia, for instance, cyberattackers laid siege to the banking, media, and other infrastructures that nearly crippled the country [11]. To protect and defend themselves from cyberattacks, these countries are increasingly outlining their position on cyberspace, cybercrime, and cybersecurity [12]. The UK Government, for instance, dedicated

£1.9 billion over 5 years (2016–21) to fund a cybersecurity program because they see cybersecurity as one of the top priorities for their country [13]. Similarly, US government agencies, including the US Department of Homeland Security, the Federal Bureau of Investigation, and several departments of the US Armed Forces, maintain cybersecurity divisions. Despite these efforts, cybersecurity is still a formidable and ever-evolving challenge because it involves a mix of physical, software, and human systems. To understand and study this system of systems, we currently rely on physical, emulated, and simulated models.

"Physical" models are those "whose physical characteristics resemble the physical characteristics of the system being modeled" [14]. In the context of cybersecurity, they are a mix of hardware and software which are connected via a network (isolated from functioning networks and the Internet) to capture a representative system. Physical models are used to test the effect of attacks or evaluate measures of protection without affecting the real system. However, they can be very costly [15] and do not incorporate the human actors that interact with the real system.

"Emulators" act in the place of a real device as part of a representative system and are usually realized as software [16, 17]. Emulators rely on virtualization or the creation of virtual machines to represent real computers and devices. Emulators provide greater flexibility than physical models because it is easier, faster, and more cost-effective to make changes in design and scale [17, 18]. However, much like physical models, emulators do not take into account the human actor and the social contracts that govern the life of an organization or society.

"Simulation models" are purposeful abstractions of physical systems [19] to explore how the complex interrelation between the human, social, software, and hardware systems might lead to vulnerability or resilience. Simulation models provide a means for examining complex interactions and changes within the system over time, including the influence of social actors. For many domains, there are significant insights that can be gained from the use of simulation models such as ecological models that aid sustainable urban development [20], social network analysis benefits from the application of simulation to the understanding diffusion [21], and ad hoc network routing [22]. Likewise, we posit that there are many benefits of using simulations for cybersecurity research.

In this article, we focus on the application of simulation to the cybersecurity domain. In this respect, we first introduce an overview of cybersecurity ("An overview of cybersecurity" section), which provides a characterization of cybersecurity research at large. We use this characterization as a guiding framework to understand state of the art in the application of simulations in the cybersecurity domain ("An overview of simulation research efforts for cybersecurity" section). Lastly, we propose a way forward for simulation research and application in cybersecurity ("Future directions" section) and conclude ("Conclusion" section) with the summary of this review article.

An overview of cybersecurity

The landscape of cybersecurity is large, ranging from individuals to nations, and continuously evolves with new threats and countermeasures. This dynamic nature of cybersecurity makes it challenging to find an objective consensus on a definition. Definitions from a sample of sources include:

i. "The state of being safe from electronic crime and the measures taken to achieve this" [23].
ii. "The activity or process, ability or capability, or state whereby information and communications systems and the information contained therein are protected from and/or defended against damage, unauthorized use or modification, or exploitation" [24].
iii. "Cyber security is the collection of tools, policies, security concepts, security safeguards, guidelines, risk management approaches, actions, training, best practices, assurance and technologies that can be used to protect the cyber environment and organization and user's assets." [25].
iv. "Prevention of damage to, protection of, and restoration of computers, electronic communications systems, electronic communications services, wire communication, and electronic communication, including information contained therein, to ensure its availability, integrity, authentication, confidentiality, and nonrepudiation." [26].

Based on these definitions, we surmise that the goal of cybersecurity is about defending and protecting cyberspace to ensure availability, integrity, and confidentiality. In that respect, we characterize cybersecurity along three dimensions: targets, threats, and preventive measures. These dimensions were inferred from the literature review and structured as a characterization of cybersecurity (Figure 1). Putting everything together, we define cybersecurity as "the practice of protecting targets and their operations from threats, through a combination of preventive measures." The purpose of the definition and its characterization is to provide a foundational understanding of the different relevant components of cybersecurity and the areas in which simulation and modeling can aid cybersecurity.

Other characterizations, such as the taxonomy of operational cybersecurity risk [27] and the cybersecurity management taxonomy [28], address specific areas of cybersecurity and provide more detail for that area. For instance, the taxonomy of operational cybersecurity risk provides very detailed subcategories like "Deliberate => Vandalism" of the "Action of People." However, with the proposed characterization, we seek to broaden the perspective to be more inclusive and capture a variety of potential scenarios.

Our definition is compatible with state-of-the-art cybersecurity concepts. The Kill Chain [29] concept of Lockheed Martin aims to deal with advanced persistent threats (APTs), which are often

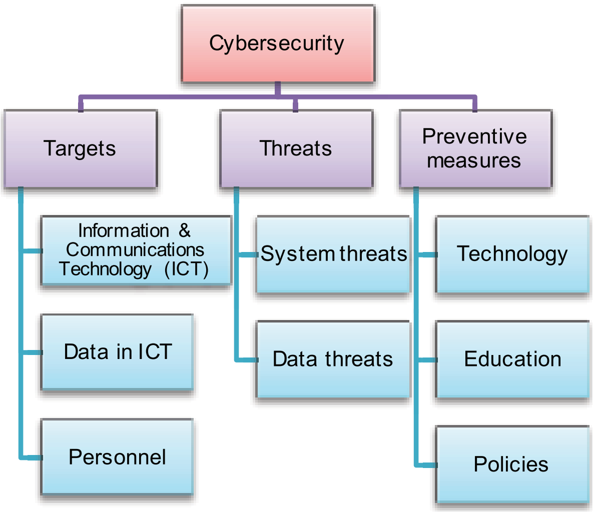

Figure 1: Characterization of cybersecurity.

conducted as a series of dependent and well-crafted actions. In comparison to our definition, "targets" are identified in the "Reconnaissance" step, while other steps from "Weaponization" to "Actions on Objectives" are captured in the form of "threats." The "preventive measures" taken to break this chain are those represented as "Courses of Actions." The Diamond Model is a well-known intrusion activity characterization model composed of four features, including "adversary," "infrastructure," "capability," and "victim" [30]. In the Diamond Model, "an adversary is the actor/organization responsible for utilizing a capability against the victim to achieve their intent" [30]. Compared to our taxonomy, "adversaries" and their "capabilities" can be represented using "threats," whereas the "infrastructure" and "victim" can be represented by "targets." Since the Diamond model is based on a cyberattack, "preventive measures" are not captured in their taxonomy. A similar comparison exists with the PrEP Framework [31], which characterizes malware attacks. The combined concept of "propagation method," "payload," and "exploits" make up the "treats on targets." Again, preventive measures are not captured in their taxonomy because their framework is based on the characterization of malware attacks. With these comparisons in place, we provide further detail on our three main pillars of cybersecurity in the following sections.

Targets

Targets refer to systems, data, and personnel of interests whose breach or access can provide benefits to nonlegitimate users or parties. These targets are categorized as Information and Communications Technology (ICT), data systems, and human systems (i.e., personnel). ICT describes the physical and networked systems that have common denominators like computing power, information processing, and computer networks that provide us the means of accomplishing tasks with greater convenience through networked infrastructures. These physical systems are often the focus of cyberattackers. Telecommunication attacks, including distributed denial of service (DDoS), routing attacks, and physical sabotage, to name a few, are ever-increasing [32]. Attacks have also been perpetrated against "secure" networks—private computer networks not connected to the Internet or telecommunication main infrastructure. The case of JPMorgan Chase is perhaps one of the most alarming ones. Peripheral devices were used to initiate the attacks; attackers used, and even infected, automated teller machine (ATM) and point-of-sale (POS) devices and were able to get into the bank's system and operate undetected for about a month [33].

Sometimes, however, the goal is to access data more than the physical system itself. Data systems or databases contain confidential information such as financial data, personal information, and information relating to national security. While increased computing power has provided a mechanism for vast amounts of data storage, it has also created a target-rich environment for cyberattacks. The attack on JPMorgan Chase's infrastructure was a lead into their data systems. It ended up relinquishing 76 million customer records with unknown secrets into the attackers' hands in addition to severe damage to the company's reputation [33]. In the information age, knowledge is power, and knowledge is generally stored and transferred through data systems. During Georgia's conflict with Russia, Internet traffic to Georgia was rerouted through Russia and Turkey for "data sniffing" [34]. These examples are just the tip of the iceberg, as of July 2019, the Identity Theft Resource Center (ITRC) identified over 10 000 publicly-noticed data breaches in the US alone, exposing over 1.6 billion data records collectively [35].

Targeting systems or data requires overcoming organizations' defense mechanisms such as firewalls and intrusion detection systems (IDSs). This is often time-consuming and expensive. For this reason, many attackers follow a way around by targeting "system users" instead. A report from IBM reveals that large percentages of cyber breaches involve some form of human error [36]. Similar findings have been found in reports published by Verizon and Symantec [37]. Social engineering, tricking users via deceptive means to obtain data or access to a system, is a common way of targeting people, especially those who work in sectors critical to a nation. Back in 2011, attackers gained access to RSA Security's servers with the use of the spear-phishing technique. While only one user was tricked into opening an Excel spreadsheet attachment with an embedded file that exploited vulnerability, it was sufficient for attackers to access the system and export sensitive data to third-party servers [38]. A similar incident occurred in Oak Ridge National Laboratory that yielded a loss of data [39]. It is important to note that RSA Security's expertise is cybersecurity, and Oak Ridge National Laboratory is one of the leading research laboratories of the US Department of Energy.

Ultimately, targeting systems, data, and people have a purpose. This purpose can be of using or selling information, modifying systems, or simply an exercise of cyber skills. The potential negative impact of such attacks is called a threat.

Threats

Threats refer to "any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, modification of information, and denial of service" [40]. We refer to attacks as the realization of threats. "Types of attacks" encompass a wide range of possibilities that rely on a simple to very sophisticated tools and algorithms focusing on system and data. For example, unpatched or 0-day (i.e., previously unknown) vulnerabilities provide access to a known flaw in software that can reveal a stepping stone for attackers with access and make the system susceptible to negative outcomes; or a socially engineered attack such as a phishing email with malware attached can relinquish system control to an attacker [41]. In an incident, US-based insurance company Anthem had millions of customer records stolen due to a vulnerability that opened the door to intruders [42]. More sophisticated attackers, once they gain access, may modify the system for further actions based on actual purpose. If the purpose of the attacker, for instance, is to access other connected systems requiring high-level permissions, the attacker likely will try escalating the privileges [43].

Another method of attacking a system is through wide-scale disruption. One of the most common ways to disrupt an Internet-connected service is to bombard it with automated calls to exhaust the service so it cannot serve actual users. The DDoS attack is a coordinated, typically large-scale attack intended to overload a system making resources unavailable to legitimate users. "Operation Ababil" is an example DDoS attack that attempted to disrupt some US-based banking companies [44]. A virus can cause a similar halt to a system as happened in the South Korean banking system [45]. And while rare, an attacker may attempt to disrupt the system by damaging or destroying it. Stuxnet is a prime example of an attack that compromised and nearly destroyed a nuclear facility in Iran [46].

Preventive measures

The best cybersecurity defense is the one that stops attacks from ever occurring. It is almost impossible to achieve as long as systems remain connected to other systems via networks or the Internet. Therefore, we rely on preventive measures which we categorize into three areas: technology, education, and policy.

"Technology" encompasses the tools, techniques, and software that detect, prevent, or stop an attack. A few of the common technologies include anti-virus software, firewalls, automated updates, and IDSs. Antivirus software recognizes known malicious programs (e.g., computer viruses, worms, Trojans) using various techniques such as pattern-based detection [47] and prevents their execution. Automated updates ensure that systems have the most up-to-date security to eliminate known vulnerabilities. IDS deals with monitoring computer and network events and analyzing them for detection of known incidents such as policy or security violations and violation attempts [48]. Techniques (i.e., protocols and algorithms) also play a vital role in preventing cyberattacks or deterring attackers. Authorization mechanisms, for instance, are a common way to secure against unauthorized use. Secure communication technologies may prevent exposure of data transferred over some networking infrastructure. In these cases, protocols like Transport Layer Security [49] and Secure Shell [50] provide a means for encrypting and protecting transferred data. Another technique for securing systems requires setting traps for would-be intruders. Honeypots, resources that have no inherent value except the ability to track and gather intelligence about attacks [51], can slow attacks, provide information about new types of attacks, and notify analysts when a system is under attack.

Technology is a prominent part of securing cyberspace. However, technology alone does not suffice. Over 90% of security incidents list human error as a factor [36]. Prevention measures, to be effective, must address the human component. "Education" and "training" are vital to cybersecurity. Training ranges from teaching users basic security concepts like safe browsing, recognizing suspicious (phishing) messages, password security, understanding software permissions, and secure data disposal, to teaching security professionals how to recognize and react to a cyberthreat.

Organizations must also be educated in cybersecurity. This occurs through the implementation and enforcement of "policy and procedures." With cyberspace being such a critical component of almost all organizations, it is necessary to describe acceptable uses and responsibilities, explicitly. Documented best practices and formal policies shared throughout organizations can aid users and improve security. Additionally, governments are crafting laws and determining enforcement protocols for cyberattacks. This, however, is beyond the scope of this article.

Synthesizing this section, we surmise that the cybersecurity domain is a combination of technical efforts that focus on protecting targets from threats through a mixture of preventive measures. Efforts to develop physical, emulation, and simulation capabilities have taken place, each with different levels of adaption [52]. Physical and emulation approaches are by far the most adopted approaches due to the realism that they bring to activities such as experimentation and testing. Research with physical systems represents the highest fidelity possible. However, testing cyberattacks on the Internet can have severe consequences. Alternatively, building a duplicate system for testing is either unfeasible or costly. Emulations allow researchers and practitioners to create virtual networks and testbeds on which they can experiment under "specific" conditions about system and network. Simulations, on the other hand, allow researchers and practitioners to test an abstraction of the system that contains only features of interest, without the need for detail, toward answering a research question. Simulations are advantageous, generally requiring less computational resources compared to emulation and physical solutions, making it cost-effective [15] and easier to scale to a large number in network size [53].

It is important to note that the literature provides perspectives of cybersecurity from two viewpoints: defending targets against threats through a combination of preventative measures and devising attacks aimed at various targets using a multitude of threats that bypass preventative measures. While it is equally important to consider both perspectives in understanding security, this work focuses on the former view.

An overview of simulation research efforts for cybersecurity

As a research area, simulation is an interdisciplinary endeavor with a vast literature. Cybersecurity research is also interdisciplinary and its literature is even larger and, making it challenging for us to truthfully capture the intersection of the two areas. To this end, we leveraged recent cybersecurity reports, our anecdotal experience from cybersecurity simulation research and development, and discussions with experts led us to cover five areas in our review. These are (i) representative environment building; (ii) test, evaluate, and explore; (iii) training and exercises; (iv) risk analysis and assessment; and (v) examining the role of people in the cybersecurity domain. There is a strong connection between these areas of research. Figure 2 shows the relationship of components that emerged from our review of simulation for cybersecurity. An operational environment, in the lexicon of Damodaran and Couretas [54], is the targetable tool in a simulation event. Operational environment building is foundational to aiding our understanding of cybersecurity and providing the necessary environment, including network topology and structure for the next research area: testing, evaluation, and exploration. These two pieces facilitate the ability to ask "what if" questions based on the operational environment. Another goal of simulation for cybersecurity is to aid in analyzing and assessing the overall risk of the system and providing enhanced training capabilities and conducting exercises. Finally, human action is introduced in cybersecurity simulation to help understand the strengths and vulnerabilities that users, attackers, and defenders bring to cybersecurity.

Representative environment building

Representative environment building refers to the creation of networks and connected systems. Research in cybersecurity requires a

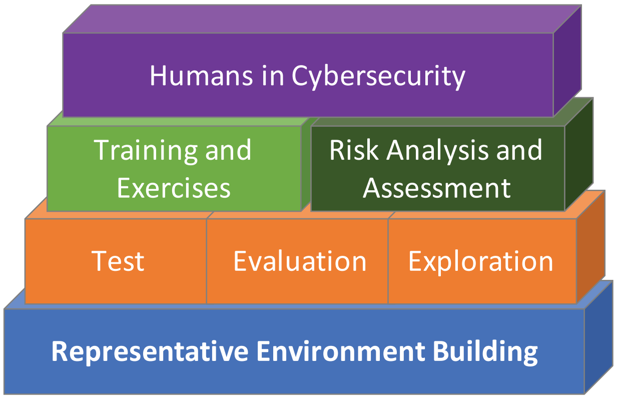

Figure 2: Identified areas of research for simulation of cybersecurity.

platform on which to test. Software-based network simulators and network traffic algorithms can be used to test specific types of cyberattacks. Open source and commercial network simulation libraries and tools are used for implementing network environment of simulations. Early attempts of network simulators date back to the 1980s and were mainly developed to test different routing and scheduling algorithms [55]. The use of such simulated networks was later possible when computing power increased and cyberattacks become prevalent. OMNeT++ [56] and INET Framework for OMNET++ [57] are robust open-source software able to simulate a vast array of devices such as wired/wireless communication networks, sensor networks, and on-chip networks, to name a few. While not as comprehensive as OMNeT++, ns-3 [58] is another popular network simulator that supports wired/wireless networks and virtualization. ns-2 [59] is still in use as of this writing although it was preceded by ns-3. There are many earlier open-source network simulation efforts such as SSFNet [60], GTNetS [61], and JiST/SWANS [62] which are not actively being developed or maintained. There are many other options for researchers to build a simulation-based representative environment. Sarkar and Halim [63] reviewed existing network simulators classified and compared based on type, deployment mode, network impairments, and protocols supported. Ojie and Pereira [64] provide a more recent review but focused on the simulation of the Internet of Things.

Network simulators usually involve only representation of devices, technologies, and communication between them while missing some critical components of a cyber scenario [52]. Others have proposed to adopt a Live, Virtual, Constructive (LVC) approach to representative environment building [54, 65, 66]. According to Damodaran and Couretas [54]:

- Live (cyber) simulation: Real actors interact with physical systems of real computers connected to real, and usually isolated networks.
- Virtual (cyber) simulation: Real actors interact with emulation or simulation of networks or emulated or simulated actors interact with real and usually isolated networks.
- Constructive (cyber) simulation: Simulated or emulated actors interact with emulations or simulations of networks.

Varshney et al. [65] developed an LVC framework called StealthNet to support testing, evaluation, and exercises using cybersecurity scenarios. Their framework involves user behavior models representing blue and red forces. According to an example scenario they present, red forces execute timed predefined actions in a DDoS scenario. Therefore, it can be used to replicate a scenario, but it cannot adapt to different evolving and dynamic conditions due to its scripted nature. Emulytics, is an LVC platform from Sandia National Lab that supports cyber training and testing [15]. Some of Emulytics capabilities include mechanisms for rapid specifications and deployment of networks, protocol support for networked devices, instantiations of networks with large numbers of nodes, and representation of wired and wireless communications. Overall, because of the number of open-source options, simulation has allowed more researchers and more areas of research to be positively impacted. One such area is security testing, evaluation, and exploration.

Test, evaluation, and exploration

The ability to explore, test, and evaluate a situation is likely the most widely used capability of simulation because of the rapid experimenting flexibility it provides. In one of the earlier examples, Zhou and Lang [67] evaluated an intrusion detection algorithm using the OPNET network simulation environment (OPNET is no longer available). To add more, Hancock and Lamont [68] examined patterns associated with intrusion detection and aided in the classification of network attacks. However, when the network gets large, it becomes a challenging task to place the IDS. Puzis et al. [69] identified the optimal placement of intrusion detection and prevention systems in such large networks. Wagner et al. [70] explored partitioning the network into sections to complicate gaining attackers access to the network. Bahşi et al. [71] analyzed literature to find out the methods employed in assessing cyber impact on missions and concluded that simulation is the dominant approach.

Cohen [72] pioneered early efforts in the simulation of cyberattacks, defenses, and consequences. While Cohen's cause and effect model was too simplistic for practical application, his efforts spurred the work of others. Chi et al. [73] continued the effort to employ simulation using a discrete-event system specification (DEVS) model with a knowledge-based learning system for the attacker and a statistical analyzer for vulnerability assessment. The simulation can classify threats, specify attack mechanisms, verify protection mechanisms, and evaluate consequences [73].

Cho and Ben-Asher [74] developed "Defense in Breadth," which is an integrated defense mechanism that combines IDS, deception systems, and moving target defense (MTD) systems. They implemented stochastic Petri nets using the Stochastic Petri Net Package (SPNP) to evaluate "the performance of a system integrated with different defense mechanisms" (p. 6). Deception systems here are mechanisms like honeypots that appear to be real parts of the network but are put in place to mislead the attacker and study its behavior. MTDs, as mentioned, enable defenders to change system behaviors, policies, or configurations automatically such that potential attack surfaces are moved dynamically.

In terms of evaluating the effect of threats and attacks, Kotenko and Ulanov [75] designed and developed an agent-based simulation framework called distributed denial of service simulator (DDoSSIM), which uses OMNeT++ and INET Framework as the network simulation basis, to evaluate different DDoS attack and defense mechanisms formed by software agents. Almajali et al. [76] utilized ns-2 [59] to create a network of a power grid and analyze the resiliency of the grid in case of DoS attack on the network.

Sonchack and Aviv [77] reported on large-scale evaluation of security systems (LESS), which simulates large-scale attacks by automatically configuring host agents based on background traffic samples and current malicious traffic models. It has the capacity to simulate 100 000 hosts with up to 5000 malicious hosts, thus providing an additional level of detail in security evaluation. Cyber Analysis Modeling Evaluation for Operations (CAMEO) [78] modeled and simulated various dynamics of a cyberecosystem such as threats and defensive strategies. CAMEO, as the authors claim, can be used to "verify and validate possible configurations and behavior for cyber agile and resilient defenses, to study sensitivity to initial configurations and discover unanticipated emergent behavior" (p. 2). Hahn and Govindarasu [79] proposed a metric to evaluate the completeness of the implemented security mechanisms in smart grids and also evaluate the applicability of the metric using simulation. It is also noted that these efforts are based on measuring traffic data. Technical systems that help recognize an attack and evaluate the possible defenses are only effective if the appropriate responses are executed. One aspect that is less explored in simulation for cybersecurity is finding conditions under which security might be vulnerable. In this case, a simulation could find a combination of input factors that lead to a desired (or undesired) output.

IDSs are usually the first line of defense in an institutional setting but they are not the only defense mechanisms to defend against malware or other malicious software. Modeling and simulation efforts support such defense approaches. Garetto et al. [80] used a special type of Markov chain to model the behavior of malware based on its activity patterns in a network. Such an approach initiates new mechanisms to catch malware other than signature-based approaches. Similar malware dynamics have been captured in agent-based and network simulation models to investigate the propagation on scale-free networks [81], Wi-Fi router networks [82], and smart grid networks [83]. Here, simulation studies allow safe testing of real or hypothetical malware spread and evaluate potential mitigation measures.

There have been recent simulation efforts that generated datasets for validation of a wide range of detection systems. Gore, Padilla [84] used Markov Chains to generate feasible APT scenarios using the Structured Threat Information eXpression language (STIX) and seed data [85]. These plausible synthetic data provide researchers with the ability to explore and understand scenarios that have not yet occurred and validate against existing detection systems as well as conduct training exercises and risk analysis. Lately, such synthetic data generation has relied on via General Adversarial Networks too. For instance, Lu and Li [86] and Kucuk and Yan [87] have developed systems to generate synthetic malware samples to be tested by a classifier. We believe there is an open venue for such simulation, especially critical domains such as the cybersecurity studies of connected vehicles because real-world data of such type is not always feasible to capture [88].

Training and exercises

Simulation can be effectively used in understanding an attack and training for different scenarios. To this end, special units have been established to conduct cybersecurity training and exercises at the national and institutional levels. For instance, the US Department of Homeland Security established the National Cyber Exercise and Planning program to support cybersecurity response plans based on strategic exercises [89]. UK's National Cyber Security Centre has similar responsibilities [90]. The NATO Cooperative Cyber Defence Centre of Excellence is responsible for organizing joint exercises regarding both technical and strategic aspects of cybersecurity [91].

There are many notable training exercises that involve a form of simulation organized by institutions listed above. A tabletop cyber training exercise called Waking Shark II drew approximately 220 participants from UK Government agencies, banks, and financial institutions. In the exercise, they simulated a disruption in the wholesale market with the goal of understanding the impact of cyberattacks and exercise communication flows between firms. The US Department of Homeland Security has conducted a similar but more comprehensive exercise called Cyber Storm since 2006. According to the official report, the latest Cyber Storm event (Cyber Storm VI) hosted a wide range of participants from "federal, state, local, tribal and territorial entities and the private sector" [92]. The goals of the exercise are reported as cybersecurity preparedness and response in order to develop and revise plans and procedures.

These large-scale, hypothetical scenarios provide a positive learning experience for the participants [93]. However, it is often difficult to convene the number of people necessary, and such large-scale gatherings are not suitable for many small to medium-size businesses. Self-paced exercises and tools are crucial to fill this gap. The UK's "Exercise in a Box" tool is a great example to satisfy this need. It focuses on organizational practices regarding a hypothetical cyber incident and response.

An alternative technique used for training is a video game-based simulation. Current cyber training games include CyberProtect used by the Department of Defense and the CyberCIEGE used by the US Navy [94]. Such game-based trainings focus on information assurance and understanding of cause-effect dynamics. There are also systems that generate learning through competition. UC Santa Barbara host the International Capture the Flag (iCFT) competition and SANS Institute host NetWars [95] that provide researchers an opportunity to learn about attack behaviors in a safe environment. CyberNEXS is used to facilitate cyberdefense competitions [95]. iCFT is geared toward university students, although it is not limited to them. NetWars and CyberNEXS engage high school students. Training and engaging the widest possible audience in cyberawareness and security are the objectives of the US National Initiative for Cybersecurity Education [96]. Simulations and competitions support such significant initiatives.

Risk analysis and assessment

As previously stated, the goal of cybersecurity is defending and protecting cyberspace through preventive measures. Creating a representative operational environment for testing, evaluating, and training against potential attacks and strategic defenses is an essential element. However, the size and scope of potential cyberattacks make total security impossible. Systems still are exposed to risks that can be analyzed and mitigated.

Risk analysis is a function that examines the likelihood of a negative outcome. When applying this concept to cybersecurity, threats are often measured regarding (i) the probability of a type of attack, (ii) the probability of attack success, and (iii) loss associated with a successful attack [97–99]. However, quantifying a loss is not a simple task in cyberspace [100]. Losses are time-dependent, not surfacing until a breach is discovered, and can reduce future value such as the loss of intellectual property. Losses can be third-party service provider-dependent—that is, interdependencies can create a loss for an entity that was not directly attacked. There are several approaches to risk analysis: probability risk assessment (PRA), attack tree analysis (ATA), fault tree analysis (FTA), and failure mode effect analysis (FMEA).

PRA quantifies risks based on statistical probabilities. PRA has general stages: identify, quantify, evaluate, and accept [101, 102]. PRA is a theoretically sound approach and has been applied in different scenarios, such as DDoS attacks against a distance learning system [103] and power grid generation losses. Monte Carlo simulations are often applied to approximate the loss of value. Despite these merits, PRA still suffers from several challenges when applied to cybersecurity: (i) historical databases are not maintained; (ii) despite a significant number of breaches, security data are not commonly shared; and (iii) the existing data is difficult to analyze for large, complex networks [102]. To this end, there are a vast number of efforts to encourage and incentivize cyber intelligence data standardization, use, and share [104, 105]. For many organizations, the decision to participate in cyber intelligence sharing programs become a perfect case for game theoretic modeling efforts. Researchers have applied both standard and evolutionary game theoretic frameworks in this domain to improve the availability of data for researchers [106, 107].

Attack trees provide a formal means of describing and analyzing the security of systems based on varying attacks [108]. Simulations have been used in both the generation and evaluation of attack trees.

The attack tree generation involves modeling of the attacker and running it to simulate attacks, which are used to generate the tree [109]. The evaluation of attack trees involves representing the system using a suitable technique and conduct Monte Carlo runs [110]. FTA uses systematic backward reasoning to determine probabilities, and FMEA uses a forward-inductive approach to do the same. These methods only examine a single fault [111].

More recently, researchers have used simulation to analyze risk across complex and interdependent systems. Charitoudi and Blyth [112] used agent-based modeling and simulation (ABMS) to estimate the cascading effects of an attack on a supply chain. Rybnicek et al. [113] combined Game Theory and ABMS to study the impact that an attack (and the subsequent defense) on critical infrastructure. Wang *et al.* [114] presented a simulation environment for analyzing and assessing the security Supervisory Control and Data Acquisition (SCADA) systems. The environment allows researchers to model vulnerability exploitation to determine the scope of the effect. Musman and Turner [115] followed a game theoretic approach to cybersecurity risk management with a cybersecurity game (CSG). The CSG uses models to describe the system, threat environment, and defender capabilities. The holistic approach is designed "for system-level analysis to inform decision-makers of good security design principles, targeted improvements, cost-effective risk reduction investments, and where defenses should be deployed" (p. 5). The types and probability of attacks, as well as the cascading effect of the interdependencies, make assessing risk in cybersecurity difficult. This is further complicated by the human factor [116]. Whether it is a careless user, a determined attacker, a skilled defender, or an insider, people are central to the use and misuse of cyberspace.

Humans in cybersecurity

Attackers, cybersecurity analysts (CSA), system administrators, and general system users interact to shape cyberspace today. Therefore, each must be considered when studying cybersecurity. Attackers can be script kiddies, state hackers, organized crime groups, insider attacker, hobbyist, hacktivist, legitimate penetration testers, or terrorist. Their role in cyberspace is defined by their skill, knowledge, resources, access, and motives or SKRAM [117]. Technology has improved the defense of cyber systems; however, the defense is still heavily dependent on who takes care of the system and who has access to it.

Human actions as they relate to simulation for cybersecurity have been far less explored compared to cyber systems especially when considering criminal behavior. Several simulations have been designed to explore network intrusion and other forms of cyberattacks. Kotenko [118], for example, modeled a DDoS attack, and Razak et al. [119] simulated network intrusions. However, these simulations do not specifically portray the attacker. Early models contained pre-scripted, static patterns, for attacker agents to follow [120]. These models eventually gave way to game theoretic and cognitive models [121], which provide a useful characterization of the initiation of an attack but ignore a host of other social contexts—user interactions, risk tolerance, social learning to name a few. Even more recent network segmentation models such as AVAIL [122] that improve on security simulations by evolving both attacker and defender strategies lack a dynamic behavioral component. Schultz [123], by contrast, examines indicators such as deliberate markers, meaningful errors, preparatory behavior, correlated usage patterns, verbal behavior, and personality traits to predict who an active insider threat might be. Along similar lines, Vernon-Bido et al. [124] investigated factors (group size, attack success rate, and opportunity) that turn a predisposed user into a cyberattacker. Moreover, Paternoster and Simpson [125], Nagin and Paternoster [126], and Hu et al. [127], used the rational choice theory to model committing a crime. These last few examples are not simulation models but can also help determining the threats that they may cause.

While attackers attempt to find unauthorized ways to use the system, CSAs must evaluate information provided by the systems to determine the level of potential risk. The risk message itself, the means of communication, and the individual decision-making process all influence the reaction to risk and alerts [128]. There is only a limited number of models on CSA behavior like Rajivan et al. [129] who developed an agent-based model to capture the cooperation of CSAs to share knowledge of attack detection. The model simulates the effectiveness of collaboration in increasing the number of alerts resolved.

Representing users in cybersecurity simulation is not widely researched. The work of Pussep et al. [130] and Blythe et al. [131] is just a couple of reported research. Blythe et al. [131], for instance, simulate system users as BDI (belief, desire, intention) agents accomplishing their routine tasks and communicating with each other. When a cyberattack is simulated, the study captures user resiliency, changes in the communication patterns, and how tasks are affected. Again, this cognitive approach captures the human as a rational decision-maker while leaving out behavioral complexities that may be more telling when considering situations like attacks based on social engineering [52].

A summary

The studies reviewed in this section are by no means exhaustive; rather they are a sample of the breadth of simulation for cybersecurity. Table 1 summarizes many of the studies presented throughout this section and highlights how they relate to the target, threat, and preventive measure characterization proposed in "An overview of cybersecurity" section. The table provides a quick reference for studies by purpose, coverage within the taxonomical representation, and simulation type.

The "purpose" column shows what the simulation is built for based on the five identified review topics. "Coverage in cybersecurity" column shows what areas of cybersecurity landscape (threats, targets, and preventive measures) these studies cover. "Type" column indicates the nature of the simulation type whether it is a platform, model, or others. Other, in this case, refers to simulation exercises that combine platform, scenarios, and models. "Names/References" column indicates citation information for studies.

Table 1 demonstrates the level of work that exists in simulation for cybersecurity. However, it also hints at the lack of focused work in certain areas. The experience of studies involving people provides insight into previously unimagined behavior by attackers and defenders. From this, we can infer that our risk analysis might also be upended when measuring the role individuals play in an organization. Models developed at the Computer Emergency Response Team at Carnegie Mellon [133], for instance, provide insight into what factors and dynamics that may lead individuals to become insiders. As such, these types of models provide evaluation and exploration capabilities for considering preventive measures. Similarly, Moskal et al. [134] considered attacker types to assess cyber threats through simulation. They examine network vulnerabilities and system configurations but also incorporate changes in the threat level of the network when the attacker is added. Adding the behavioral

Table 1: Summary of the surveyed papers according to their purpose, coverage in cybersecurity, and type

Purpose					Coverage in cybersecurity			Type	Name (reference)
Representative env. building	Test, evaluation, and exploration	Training and exercises	Risk analysis and assessment	Humans in cybersecurity	Targets	Threats	Preventive measures		
✓					✓			Platform	• OMNeT++ [56] with INET Framework [57] • ns-3 [58] • ns-2 [59] • SSFNet [60] • GTNetS [61] • JiST/SWANS [62]
✓	✓				✓	✓	✓	Platform	• LESS [77] • CAMEO [78] • Kotenko and Ulanov [75] • Kotenko [118]
✓	✓	✓			✓	✓	✓	Platform	• StealthNet [65]
✓	✓	✓			✓			Platform	• Emulytics [15]
		✓				✓	✓	Other	• Walking Shark II [93] • Cyber Storm [92]
		✓	✓		✓	✓	✓	Platform	• CyberProtect; CyberCIEGE [94]; Exercise in a Box
	✓				✓	✓	✓	Model	• Zhou and Lang [67] • Hancock and Lamont [68] • Puzis et al. [69] • Wagner et al. [70] • Cohen [72] • Chi, Park [73] • Cho and Ben-Asher [74] • Almajali et al. [76] • Hahn and Govindarasu [79] • Razak et al. [119]
	✓					✓	✓	Model	• Garetto et al. [80] • Hosseini et al. [81] • Eder-Neuhauser et al. [83]
	✓	✓	✓			✓		Model	• Gore, Padilla [84] • Lu and Li [86] • Kucuk and Yan [87]
	✓		✓		✓	✓		Model	• Kavak et al. [82] • Keskin et al. [103] • Tatar et al. [132]
			✓				✓	Model	• Tosh et al. [106, 107]
			✓			✓		Model	• Karray et al. [109]
			✓			✓	✓	Model	• Dalton et al. [110]
	✓			✓	✓	✓	✓	Model	• Hamilton and Hamilton [120] • Dutt et al. [121]
	✓		✓		✓	✓	✓	Model	• Charitoudi and Blyth [112] • Rybnicek et al. [113] • Musman and Turner [115]
	✓		✓		✓	✓	✓	Platform	• Wang, Fang [114]
	✓			✓	✓	✓	✓	Model	• AVAIL [122]
	✓			✓	✓	✓		Model	• Vernon-Bido et al. [124] • Rajivan et al. [129] • Pussep et al. [130] • Blythe et al. [131]

component to test beds and traffic generators like the Lincoln Adaptable Real-time Information Assurance Testbed or LARIAT [135] will aid in the advancement of cybersecurity simulations.

There are drawbacks related to modeling a system of systems like cybersecurity. According to Windrum et al. [136], designing representative models for simulation has several core issues of validation: (i) models are generally abstractions that focus on relationships in isolation of the real, and often unknown, system (ii) models are designed under assumptions, some of which may change with empirical data while others are never subjected to empirical validation (iii) the number of variables used to describe the model are limited, many variables are simplified or omitted to aid in understanding a particular relationship and (iv) models represent a possible theory based on data; however, data can also support alternate theories; and testing and falsifying a hypothesis in isolation is not accurate as the systems do not operate in isolation.

Simulation for cybersecurity is found in operational environment building as well as in testing, evaluating, training, and risk analyzing phases of cybersecurity. The elements of complexity, interdependence, and social connectivity show why simulation is an effective tool for this research. However, there is still a lot more potential when it comes to the use of simulation for cybersecurity.

Future directions

Cyberspace is a sociotechnical realm, and simulation is probably the best tool available to explore this integration. Simulation provides an arena for integrating operational environments with human behavior models to analyze the vulnerability of a network structure given different types of users. Simulations give insight into the possible effects of the use and misuse of cyber systems. In an interconnected society, simulation can show how attacks to one critical infrastructure strain the entire system. As such, simulations can aid in understanding the ripple effects of cyberattacks on the system and society. We believe that to capture these dynamics effectively, there is a need for more focused efforts. Our review in “An overview of simulation research efforts for cybersecurity” section and current trends in cybersecurity reveal that many independent efforts address a piece in cybersecurity. There is a pressing need to develop synergetic efforts to improve the coverage, quality, and reuse of existing studies. In this respect, we identified three main avenues for future research to advance both cybersecurity and simulation.

i. Advance data collection and access
ii. Generate new theoretical constructs
iii. Improve behavioral models for simulation

Ouyang [137] notes one of the challenges of exploring the interconnection of critical infrastructure is the lack of precise data. The same is true for cybersecurity in general. Data is generally sparse, incomplete, or unavailable due to its sensitive nature. Similarly, there have been efforts that encourage and illustrate the usefulness of cybersecurity research datasets [138]. However, there is another problem with the data—often, the right data is not collected. Securing systems that affect the entire population requires cultural data, economic data, and political data in addition to system and threat information. Cybersecurity simulation models need data that reflects the current social climate and norms. Prevention is most effective when we can understand and control the environment that gives rise to an attack.

To this end, data collection needs to focus not just on the attack but on the environment that gives rise to the attack. Data collection methods for cybersecurity simulations should feed deep learning [139] and other artificial intelligence (AI) techniques. These tools provide a vibrant multilayered approach to representing vast amounts of data, including social awareness data, which until now might have seemed unrelated. AI and tools like Watson use cognitive computing to harness the potential of the massive data that are available [140]. They offer the possibility of finding connections in data through simulations with greater context. Current work in this area includes the emerging AI-based automated cybersecurity decision systems and their integration with simulations [70].

New data sources providing more information about the context of attacks might lead to new theoretical constructs. Current theoretical constructs in cybersecurity tend to be disparate and heavily dependent on empiricism [141]. Foundational theories of network science and cybersecurity center around attackers, defenders, networked assets, and policies [142]. The defense technology continues to advance at a rapid pace; defense against kernel rootkits [143] and block chain-based data provenance [144] advance cybersecurity into the cloud environments. However, attack strategies and maneuvers around defenses move at an even faster pace infecting the system with zombie bots and APTs. And vulnerabilities increase as the complexity of the system expands. New theoretical constructs should integrate technology advances with theories of motivation, behavioral analysis, and criminology to expand the science of cybersecurity to cover the full spectrum of this human-made universe. Tisdale [28] suggested a new construct based on the combination of Systems Theory [145] and Complexity Leadership Theory [146]. These two theories provide a strong foundation, but it is only a beginning. Theories, like routine activity theory [147]—that stipulates that crime requires three conditions: likely offender, a suitable target, and absence of capable guardian—need to be updated to consider the motivations of attackers in a world of unsuitable targets with the presence of capable guardians. Simulation models provide a theory-generation capability by modeling human behavior in cybersecurity environments. The traditional “arms race” mentality focuses internally on improving the technical defense. A defense construct, and defense models, should include an understanding of the adversary in a manner that adjusts to behaviors and cultural norms observed in the attacker.

A sociotechnical perspective forces us to rethink not only how we conduct cybersecurity, but also how we implement that perspective when considering targets, threats, and preventive measures. This becomes even more critical with the expansion of cyberphysical systems (CPS). Hybrid simulation for cyberphysical systems [148] describes the challenges associated with designing and testing CPS but does not even address the massive security challenges that these systems pose. As CPS becomes a larger part of everyday life, this true marriage of social and technical systems must incorporate new theoretical constructs and human behavior models into the simulations [149].

Cybersecurity systems and behavioral models should range from the actions of individual users through attackers and defenders and into policymakers. Open problems such as attribution call for solutions that combine data analysis and simulation. There is a need to define the level of granularity concerning adversarial strategies vs. tactics vs. intent, where state of the art is and challenges. It is important to highlight that a new perspective that considers the role of people in cybersecurity is needed.

Models of humans and organizations are needed to provide an exploratory ground for studying how social engineering campaigns, for instance, can take place or what conditions may lead users to become insider threats over the long term. Activities like risk analysis may be completely upended when measuring the role individuals play in an organization. However, modeling human behavior in cybersecurity presents several challenges, mainly due to the lack of theories from psychology, criminology, or sociology that can provide a finer level of detail in the modeling of individuals. The effectiveness of simulations for cyber will be determined by several factors, among which is the need for interdisciplinary work.

Simulation of cybersecurity, just like cybersecurity in general, has focused primarily on the technical aspects. As we incorporate users into the equation, we need to focus on the behavioral aspects. Securing cyberspace is not simply a function of better technology; to improve security, we need to understand the attacker, the user, the defender, the organizational context to better model targets, threats, and preventive measures. Psychologists and criminologists should be key collaborators in developing solutions in the cybersecurity domain.

Conclusions

The goal of this article is to provide an overview of cybersecurity, a comprehensive review of significant simulation efforts for cybersecurity, and propose a way forward for advancing both simulation and cybersecurity areas. It begins by inferring the dimensions of cybersecurity—threats, targets, and preventative measures—to provide a foundational understanding of the cybersecurity landscape. Targets refer to systems or networks of interests whose breach can provide benefit to nonlegitimate users. Threats or attacks refer to the negative outcomes of nonlegitimate access to cyber systems, data, or people due to different types of attack conducted by different vectors (attacker or attacking tool). Preventive measures refer to all efforts that attempt to reduce the probability of attacks ever occurring through technology, education, and policy.

The article reviews some of the important areas that simulation is used in cybersecurity—representative environment building; test, evaluation, and exploration; training and exercises; risk analysis and assessment; and exploring the humans in cybersecurity. Representative environment building refers to the creation of networks and connected systems consisting of physical or virtual machines with simulation providing a flexible alternative. Test, evaluation, and exploration in cybersecurity involve recognizing when an attack is occurring, understanding the effects of different attacks, and learning which responses are most effective against the attack. Training and exercises, like in other areas where simulation is used, provides a platform (less expensive and more flexible) for acquiring specific skills. Risk analysis and assessment refer to the use of simulations as the means to assess risk metrics to potentially develop policies, approaches, or technology to minimize such risk. Lastly, modeling humans in the context of cyber is a major line of research that puts people as part of cyber defense and attack. Attackers, CSA, and general system users interact to shape cyberspace as it exists today, especially when considering organizational context like costs associated with cyber disruption and loss of organizational image.

Last, we discussed a proposed way forward by developing means to establish data collection and access to inform models, to use existing social theories to create new theoretical constructs specific to the cybersecurity domain, and considering behavioral models in cyber as the means to develop sociotechnical solutions. The challenges of cybersecurity are vast and growing. Securing cyberspace is an imposing task that will take great technical ability combined with behavioral insights. We argued that simulations should play a bigger role, while more focused research is needed in this area.

Acknowledgments

The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the OASD(R&E) or the US Government.

Funding

This material is based on research sponsored by the Office of the Assistant Secretary of Defense for Research and Engineering [OASD(R&E)] under agreement number FAB750-15-2-0120.

Conflict of interest statement. Authors declare no conflict of interest.

References

1. Radack S. *Managing Information Security Risk: Organization, Mission and Information System View, in ITL Bulletin*. NIST Information Technology Laboratory - Computer Security Resource Center, 2011.
2. Maughan D. The need for a national cybersecurity research and development agenda. *Communications of the ACM* 2010;**53**:29.
3. Good VR. *Identity Theft and the Internet*. Utica College, 2019, p. 48.
4. Poyraz OI, Bouazzaoui S, Keskin O *et al.* Cyber-assets at risk (CAR): The cost of personally identifiable information data breaches. In: *International Conference on Cyber Warfare and Security*, 2020. Norfolk, Virginia, USA: Academic Conferences International Limited.
5. Sheppard D. *ID Theft down 28 Percent in U.S. in 2010: Survey*. New York: Reuters, 2011.
6. Marchini K, Pascual A. *2019 Identity Fraud Study: Fraudsters Seek New Targets and Victims Bear the Brunt*, in *2019 Identity Fraud Study*. Javelin, 2019, 45.
7. Poyraz OI, Canan M, McShane M *et al.* Cyber assets at risk: monetary impact of US personally identifiable information mega data breaches. *The Geneva Papers on Risk and Insurance-Issues and Practice* 2020;**45**:616–38.
8. Lewis J. *Economic Impact of Cybercrime – No Slowing Down*, Santa Clara, CA, USA: McAfee LLC, 2018.
9. Gorman S. Electricity grid in U.S. penetrated by spies. *The Wall Street Journal*, 2009, 3–5.
10. Thakur K, Ali ML, Jiang N *et al.* Impact of cyber-attacks on critical infrastructure. In: *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)*, 2016. New York, NY, USA.
11. Ottis R. Analysis of the 2007 cyber attacks against estonia from the information warfare perspective. In: *Proceedings of the 7th European Conference on Information Warfare*, 2008.
12. Von Solms R, Van Niekerk J. From information security to cyber security. *Comput Secur* 2013; **38**:97–102.
13. Office UC. *National Cyber Security Strategy 2016-2021*. UK Cabinet Office, 2016.
14. DoD, DoD Modeling *and* Simulation Glossary, Department of Defense. Under Secretary of Defense for Acquisition Technology, 1998.
15. Leeuwen BV, Urias V, Stout W, *et al.* Emulytics at Sandia National Laboratories. In: *MODSIM World*, 2015. Virginia Beach, VA, USA.
16. Guruprasad S, Ricci R, Lepreau J. Integrated network experimentation using simulation and emulation. In: *Proceddings - First International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, Tridentcom 2005*, 2005, 204–12.

17. Rimondini M. *Emulation of computer networks with Netkit*. 2007, Roma, Italy: Universita̓ di Roma Tre.
18. Pizzonia M, Rimondini M. Netkit: network emulation for education. Software: Practice and Experience 2016;**46**:133–165. 10.1002/spe.2273
19. Turnitsa C, Padilla, JJ Tolk A. *Ontology for modeling and simulation*. Proceedings - Winter Simulation Conference, 2010: p. 643–651.
20. Aburas MM, Ho YM, Ramli MF *et al.* The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: a review. *Int J Appl Earth Obs Geoinf* 2016;**52**:380–9.
21. Kiesling E, Günther M, Stummer C *et al.* Agent-based simulation of innovation diffusion: a review. *Cent Eur J Oper Res* 2012;**20**:183–230.
22. Stanica R, Chaput E, Beylot A-L. Simulation of vehicular ad-hoc networks: challenges, review of tools and recommendations. *Comput Netw* 2011;**55**:3179–88.
23. Dictionary CE. *In Collins English Dictionary*. HarperCollins Publishers, 2020.
24. Studies, N.I.f.C.C.a. *A Glossary of Common Cybersecurity Terminology*, 2020 November 28, 2018. https://niccs.us-cert.gov/about-niccs/glossary (20 March 2020, date last accessed).
25. ITU. Overview of cybersecurity, in Series X: *Data Networks, Open System Communications and Security*. International Telecommunication Union, 2008.
26. CNNS. Committee on National Security Systems (CNNS) Glossary. Ft Meade, MD, USA: CNNS, 2015.
27. Cebula JJ, Popeck ME, Young LR. *A Taxonomy of Operational Cyber Security Risks Version 2*. 2014. Software Engineering Institute.
28. Tisdale SM. Cybersecurity: challenges from a systems, complexity, knowledge management and business intelligence perspective. *Issues Infor Syst* 2015;**16**:191–8.
29. Hutchins EM, Cloppert MJ, Amin RM. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Iss Inform Warf Secur Res* 2011;**1**:80.
30. Caltagirone S, Pendergast A, Betz C. *The Diamond Model of Intrusion Analysis*. Center for Cyber Intelligence Analysis and Threat Research Hanover Md: Ft. Meade, MD, USA, 2013.
31. Herr T. A framework for malware & cyber weapons PrEP. *J Inform Warf* 2014;**13**:87–106.
32. Japertas S, Baksys T. Method of early staged cyber attacks detection in IT and telecommunication networks. *E. ir Elekt,* 2018;**24**:68–77.
33. Agarwal T, Henry D, Finkle J. *JPMorgan Hack Exposed Data of 83 Million, among Biggest Breaches in History, in Reuters*. Thomson Reuters, 2014.
34. Shakarian P. The 2008 Russian cyber campaign against Georgia. *Military Rev* 2011;**91**:63.
35. ITRC. *Identity Theft Resource Center*, 2020. https://www.idtheftcenter.org/ (27 March 2020, date last accessed).
36. IBM. *IBM 2015 Cyber Security Intelligence Index*, 2015, IBM.
37. Menn J. *User Mistakes Aid Most Cyber Attacks, Verizon and Symantec Studies Show*, in *Reuters*. Thomson Reuters, 2015.
38. Krombholz K, Hobel H, Huber M *et al.* Advanced social engineering attacks. *J Inform Secur Appl* 2015;**22**:113–22.
39. Laszka A, Vorobeychik Y, Koutsoukos X. Optimal personalized filtering against spear-phishing attacks. In: *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015. Austin, TX, USA: AAAI Press.
40. Kissel R. *Glossary of Key Information Security Terms*. National Institute of Standards and Technology, 2013.
41. Kucuk Y, Patil N, Shu Z *et al.* BigBing: privacy-preserving cloud-based malware classification service. In: *2018 IEEE Symposium on Privacy-Aware Computing (PAC)*, 2018. Washington DC, USA: IEEE.
42. Balbi A. Massive cyber attack at anthem. *Strat Financ* 2015;**96**:11.
43. Snyder P, Kanich C. One thing leads to another: credential based privilege escalation. In: *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy*. San Antonio, TX: Association for Computing Machinery, 2015, 135–7.
44. Gillman D, Lin Y, Maggs B *et al.* Protecting websites from attack with secure delivery networks. *Computer* 2015;**48**:26–34.
45. Sang-Hun C. Computer networks in South Korea are paralyzed in cyberattacks. *The New York Times*, 2013.
46. Langner R. Stuxnet: dissecting a cyberwarfare weapon. *IEEE Secur Priv* 2011;**9**:49–51.
47. Szor P. *The Art of Computer Virus Research and Defense*. Upple Sadle River, NJ, USA: Addison-Wesley Professional, 2005.
48. Singh R, Kumar H, Singla RK *et al.* Internet attacks and intrusion detection system: a review of the literature. *Online Inform Rev* 2017;**41**: 171–84.
49. Dierks T, Rescorla E. The Transport Layer Security (TLS) Protocol Version 1.2, *in RFC 5246*, 2008, The IETF Trust.
50. Barrett DJ, Silverman RE, Byrnes RG. *SSH, the Secure Shell: The Definitive Guide: The Definitive Guide*. O'Reilly Media, 2005.
51. Paradise A, Shabtai A, Puzis R *et al.* Creation and management of social network honeypots for detecting targeted cyber attacks. *IEEE Trans Comput Soc Syst* 2017;**4**:65–79.
52. Kavak H, Padilla JJ, Vernon-Bido *et al. A Characterization of Cybersecurity Simulation Scenarios*. Pasadena, CA: ACM, 2016.
53. Calheiros RN, Netto MAS, Buyya R. EMUSIM: an Integrated Emulation and Simulation Environment for Modeling, Evaluation, and Validation of Performance of Cloud Computing Applications. *Softw Pract Exp* 2012;**39**:1–18.
54. Damodaran SK, Couretas JM. Cyber modeling & simulation for cyber-range events. In: *Summer Computer Simulation Conference*. Chicago, IL: Society for Modeling & Simulation International (SCS), 2015.
55. Keshav S. *REAL: A Network Simulator*. University of California Berkeley, USA. 1988.
56. Varga A, Hornig R. An overview of the OMNeT++ simulation environment. In: *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering): Marseille, France. p. Article 60, 2008.
57. INET. *INET Framework - INET Framework*, 2020. https://inet.omnetpp.org/ (19 April 2020, date last accessed).
58. Henderson TR, Lacage M, Riley GF *et al.* Network simulations with the ns-3 simulator. *SIGCOMM Demonst* 2008;**14**:527.
59. Issariyakul T, Hossain E. *Introduction to Network Simulator 2 (NS2)*, in *Introduction to network simulator NS2*. New York, NY, USA: Springer, 2009, p, 19–36.
60. Yoon S, Kim YB. *A design of network simulation environment using ssfnet*. in *2009 First International Conference on Advances in System Simulation*. 2009. IEEE.
61. Riley GF. The Georgia tech network simulator. In: *Proceedings of the ACM SIGCOMM workshop on Models, methods and tools for reproducible network research*, 2003.
62. Barr R. *Swans-Scalable Wireless Ad Hoc Network Simulator.* User Guide, 2004.
63. Sarkar NI, Halim SA. A review of simulation of telecommunication networks: simulators, classification, comparison, methodologies, and recommendations. *Cyber J* 2011;**2**:10–17.
64. Ojie E, Pereira E. Simulation tools in internet of things: a review. In: *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*, 2017. Liverpool, UK: ACM.
65. Varshney M, Pickett K, Bagrodia R. A Live-Virtual-Constructive (LVC) framework for cyber operations test, evaluation and training. In: *Proceedings - IEEE Military Communications Conference MILCOM*, 2011, 1387–92.
66. Bergin DL. Cyber-attack and defense simulation framework. *J Defens Model Simul* 2015;**12**:383–92.
67. Zhou, M. and S.-d. Lang, A Frequency-based approach to intrusion detection. Systemics, Cybernetics, and Informatics, 2003. 2(3): p. 52–56.
68. Hancock DL, Lamont GB. Multi agent system for network attack classification using flow-based intrusion detection. In: *2011 IEEE Congress of Evolutionary Computation, CEC 2011*, 2011, 1535–42.
69. Puzis R, Tubi M, Elovici Y *et al.* A decision support system for placement of intrusion detection and prevention devices in large-scale networks. *ACM Trans Model Comput Simul* 2011;**22**:1–26.
70. Wagner N, Sahin CS, Winterrose M *et al.* Towards automated cyber decision support: A case study on network segmentation for security. In:

2016 IEEE Symposium Series on Computational Intelligence (SSCI), 2016.

71. Bahşi H, Udokwu CJ, Tatar U *et al.* Impact assessment of cyber actions on missions or business processes: A systematic literature review. In: *ICCWS 2018 13th International Conference on Cyber Warfare and Security*. Academic Conferences and publishing limited, 2018.
72. Cohen F. Simulating cyber attacks, defences, and consequences. *Comput Secur* 1999;**18**:479–518.
73. Chi SD, Park JS, Jung KC *et al. Network security modeling and cyber attack simulation methodology. In: Information Security and Privacy*. Berlin: Springer Berlin Heidelberg, 2001; 320–333.
74. Cho J-H, Ben-Asher N. Cyber defense in breadth: modeling and analysis of integrated defense systems. *J Def Model Simul* 2018;**15**:147–60.
75. Kotenko IV, Ulanov A. Simulation of Internet DDoS attacks and defense. In: *Proceedings of the 9th International Conference on Information Security*, 2006, 327–42.
76. Almajali A, Viswanathan A, Neuman C. Analyzing resiliency of the smart grid communication architectures under cyber attack. In: *Proceedings of 5th ACM USENIX Workshop on Cyber Security Experimentation and Test (USENIX CSET 2012)*, 2012.
77. Sonchack J, Aviv AJ. LESS Is More: Host-Agent Based Simulator for Large-Scale Evaluation of Security Systems. In: *European Symposium on Research in Computer Security*, 2014.
78. Hassell S, Beraud P, Cruz A *et al.* Evaluating network cyber resiliency methods using cyber threat, vulnerability and defense modeling and simulation. In: *Proceedings - IEEE Military Communications Conference MILCOM*, 2012 (August 2009).
79. Hahn A, Govindarasu M. Cyber attack exposure evaluation framework for the smart grid. *IEEE Trans Smart Grid* 2011;**2**:835–43.
80. Garetto M, Gong W, Towsley D. Modeling malware spreading dynamics. In: *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)*, 2003.
81. Hosseini S, Azgomi MA, Torkaman AR. Agent-based simulation of the dynamics of malware propagation in scale-free networks. *Simulation* 2016;**92**:709–22.
82. Kavak H, Padilla JJ, Vernon-Bido D *et al.* The spread of Wi-Fi router malware revisited. In: *Spring Simulation Multi-Conference*. Virginia Beach, VA: ACM, 2017.
83. Eder-Neuhauser P, Zseby T, Fabini J. Malware propagation in smart grid networks: metrics, simulation and comparison of three malware types. *J Comput Virol Hack Tech* 2019;**15**:109–25.
84. Gore R, Padilla J, Diallo S. Markov chain modeling of cyber threats. *J Def Model Simul* 2017;**14**:233–44.
85. Barnum S. *Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX)*. Mitre Corporation, 2012, **11**, 1–22.
86. Lu Y, Li J. Generative adversarial network for improving deep learning based malware classification. In: *2019 Winter Simulation Conference (WSC)*, 2019, IEEE.
87. Kucuk Y, Yan G. Deceiving portable executable malware classifiers into targeted misclassification with practical adversarial examples. In: *Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy*, 2020.
88. Rajbahadur GK, Malton AJ, Walenstein A *et al.* A survey of anomaly detection for connected vehicle cybersecurity and safety. In: *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018.
89. CISA. National Cyber Exercise And Planning Program, 2020. https://www.cisa.gov/national-cyber-exercise-and-planning-program (20 March 2020, date last accessed).
90. NCSC. National Cyber Srcurity Centre, 2020. https://www.ncsc.gov.uk/ (20 March 2020, date last accessed).
91. NATO. *CCDCOE - The NATO Cooperative Cyber Defence Centre of Excellence Is a Multinational and Interdisciplinary Hub of Cyber Defence Expertise,* 2020 https://ccdcoe.org/ (20 March 2020, date last accessed).
92. CISA, *Cyber Storm VI: National Cyber Exercise*, D.o.H. Security, Editor, 2020.
93. Keeling C. *Waking Shark II Desktop Cyber Exercise - Report to Participants*, 2013.
94. Cone BD,Thompson MF, Irvine CE *et al.* Cyber security training and awareness through game play. *IFIP Int Feder for Inform Process* 2006; **201**:431–36.
95. Nagarajan A, Allbeck JM, Sood A *et al.* Exploring game design for cybersecurity training. In: *2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2012, 256–62.
96. Furman S, Theofanos MF, Choong Y-Y *et al.* Basing cybersecurity training on user perceptions. *IEEE Secur Priv* 2012;**10**:40–9.
97. Gordon LA, Loeb MP. *Managing Cybersecurity Resources: A Cost-Benefit Analysis*. Vol. **1**. New York: McGraw-Hill, 2006.
98. Sommestad T, Ekstedt M, Johnson P. A probabilistic relational model for security risk analysis. *Comput Secur* 2010;**29**:659–79.
99. Tatar Ü, Karabacak B. An hierarchical asset valuation method for information security risk analysis. In: *International Conference on Information Society (i-Society 2012)*. IEEE, 2012.
100. Rowe C, Seif Zadeh H, Garanovich IL *et al.* Prioritizing investment in military cyber capability using risk analysis. *J Def Model Simul* 2019;**16**: 321–33.
101. Haimes Y. Risk modeling, assessment, and management. *IEEE Trans Syst Man Cybernetics C (Appl Rev)* 1999;**29**:315.
102. Taylor C, Krings A, Alves-Foss J. Risk analysis and probabilistic survivability assessment (RAPSA): an assessment approach for power substation hardening. In: *Proc. ACM Workshop on Scientific Aspects of Cyber Terrorism,(SACT), Washington DC*, 2002.
103. Keskin O *et al.* Economics-Based Risk Management of Distributed Denial of Service Attacks: A Distance Learning Case Study. In: *ICCWS 2018 13th International Conference on Cyber Warfare and Security*. Academic Conferences and publishing limited, 2018.
104. Skopik F, Settanni G, Fiedler R. A problem shared is a problem halved: a survey on the dimensions of collective cyber defense through security information sharing. *Comput Secur* 2016;**60**:154–76.
105. Vázquez DF, Acosta OP, Spirito C *et al. Conceptual framework for cyber defense information sharing within trust relationships*. In: *2012 4th International Conference on Cyber Conflict (CYCON 2012)*. IEEE, 2012.
106. Tosh D, Sengupta S, Kamhua C *et al.* An evolutionary game-theoretic framework for cyber-threat information sharing. In: *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015.
107. Tosh D, Molloy M, Sengupta S *et al.* Cyber-investment and cyber-information exchange decision modeling. In: *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*. IEEE, 2015.
108. Schneier B. Attack trees. *Dr Dobb's J* 1999;**24**:21–9.
109. Karray K, Danger JL, Guilley S *et al. Attack tree construction and its application to the connected vehicle. In: Cyber-Physical Systems Security*. Springer, 2018, 175–90.
110. Dalton G, Mills RF, Colombi JM *et al.* Analyzing attack trees using generalized stochastic petri nets. In: *Information Assurance Workshop*. IEEE, 2006.
111. Ralston PA, Graham JH, Hieb JL. Cyber security risk assessment for SCADA and DCS networks. *ISA Trans* 2007;**46**:583–94.
112. Charitoudi K, Blyth A. An agent-based socio-technical approach to impact assessment for cyber defense. In: *Proceedings - 4th International Conference on Emerging Intelligent Data and Web Technologies, EIDWT 2013*, 2013, 558–63.
113. Rybnicek M, Tjoa S, Poisel R. Simulation-based cyber-attack assessment of critical infrastructures marlies. *Lect Notes Bus Inf Process* 2014;**191**: 135–50.
114. Wang C, Fang L, Dai Y. A simulation environment for SCADA security analysis and assessment. In: *2010 International Conference on Measuring Technology and Mechatronics Automation*. IEEE, 2010.
115. Musman S, Turner A. A game theoretic approach to cyber security risk management. *J Def Model Simul* 2018;**15**:127–46.

116. Vernon-Bido D, Grigoryan G, Kavak H *et al.* Assessing the impact of cyberloafing on cyber risk. In: *Proceedings of the Annual Simulation Symposium*, 2018.
117. Parker DB. *Fighting Computer Crime: A New Framework for Protecting Information*. New York, NY, USA: John Wiley & Sons, Inc, 1998.
118. Kotenko IV. Multi-agent modelling and simulation of cyber-attacks and cyber-defense for homeland security. In: 2007 *4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS*, 2007(September), 614–9.
119. Razak S, Zhou M, Lang S-D. Network intrusion simulation using OPNET. In: *OPNETWORK2002 Conference*. Citeseer, 2002.
120. Hamilton SN, Hamilton WL. Adversary modeling and simulation in cyber warfare. *IFIP Int Feder Inf Process* 2008;**278**:461–75.
121. Dutt V, Ahn Y-S, Gonzalez C. Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. *Hum Fact* 2013;**55**:605–18.
122. Hemberg E, Zipkin JR, Skowyra RW *et al.* Adversarial co-evolution of attack and defense in a segmented computer network environment. In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2018.
123. Schultz EE. A framework for understanding and predicting insider attacks. *Comput Secur* 2002;**21**:526–31.
124. Vernon-Bido D, Padilla JJ, Diallo S *et al. Towards Modeling Factors That Enable an Attacker*. Montreal, QC: Society for Modeling & Simulation International (SCS), 2016.
125. Paternoster R, Simpson S. Sanction threats and appeals to morality: testing a rational choice model of corporate crime. *Law and Society Review* 1996;**30**:549–83.
126. Nagin DS, Paternoster R. Enduring individual differences and rational choice theories of crime. *Law Soc Rev* 1993;**27**:467–96.
127. Hu Q, Xu Z, Dinev T *et al.* Does deterrence work in reducing information security policy abuse by employees? *Commun ACM* 2011;**54**: 54–60.
128. Nurse JR, Creese S, Goldsmith M *et al.* Trustworthy and effective communication of cybersecurity risks: a review. In: *2011 1st Workshop on Socio-Technical Aspects in Security and Trust (STAST)*. IEEE, 2011.
129. Rajivan P, Janssen MA, Cooke NJ. Agent-based model of a cyber security defense analyst team. In: *Proceedings of the human factors and ergonomics society annual meeting*. Los Angeles, CA: SAGE Publications Sage CA, 2013.
130. Pussep K, Leng C, Kaune S. Modeling user behavior in p2p systems. In K. Wehrle, G. M, and J. Gross, eds. *Modeling and Tools for Network Simulation*, **2010**, Berlin, Heidelberg: Springer, p. 447-461.
131. Blythe J, Botello A, Sutton J *et al. Testing Cyber Security with Simulated Humans*. Artificial Intelligence, 2011, 1622–7.
132. Tatar U, Bahsi H, Gheorghe A. Impact assessment of cyber attacks: a quantification study on power generation systems. In *2016 11th System of Systems Engineering Conference (SoSE)*. IEEE, 2016.
133. Moore AP, Cappelli DM, Trzeciak RF. The “big picture” of insider IT sabotage across US critical infrastructures. In: *Insider Attack and Cyber Security*, S. Stolfo, et al., Editors. 2008, Springer, Boston, MA. 17-52.
134. Moskal S, Yang SJ, Kuhl ME. Cyber threat assessment via attack scenario simulation using an integrated adversary and network modeling approach. *J Def Model Simul* 2018;**15**:13–29.
135. Haines JW , Rossey LM, Lippmann RP *et al.* Extending the darpa offline intrusion detection evaluations. In: *Proceedings DARPA Information Survivability Conference and Exposition II. DISCEX’01*. IEEE, 2001.
136. Windrum P, Fagiolo G, Moneta A. Empirical validation of agent-based models: alternatives and prospects. *J Artif Soc Soc Simul* 2007;**10**.
137. Ouyang M. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliab Eng Syst Safe* 2014;**121**:43–60.
138. Moore T, Kenneally E, Collett M *et al.* Valuing Cybersecurity Research Datasets. In: Moore T, Kenneally E, Collett M, Thapa P. *Valuing Cybersecurity Research Datasets*. In: *18th Workshop on the Economics of Information Security (WEIS)*, 2019.
139. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**: 436–44.
140. Chen Y, Argentinis JE, Weber G. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research. *Clin Ther* 2016;**38**:688–701.
141. McMorrow D. *Science of Cyber-Security*. Mitre Corp: McLean, VA, USA, 2010.
142. Kott A. Towards fundamental science of cyber security. In R. Pino ed. *Network Science and Cybersecurity* 2014, Springer: New York, NY. 1-13.
143. Zhang L , Shetty S, Liu P *et al.* Rootkitdet: Practical end-to-end defense against kernel rootkits in a cloud environment. In: *European Symposium on Research in Computer Security*. Wroclaw, Poland: Springer, 2014.
144. Liang X, Shetty S, Tosh D *et al.* Provchain: a blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. In: *2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*. IEEE, 2017.
145. Von Bertalanffy L. General system theory. *Gen Syst* 1956;**1**:11–7.
146. Uhl-Bien M, Marion R, McKelvey B. Complexity leadership theory: shifting leadership from the industrial age to the knowledge era. *Leadersh Q* 2007;**18**:298–318.
147. Cohen LE, Felson M. Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 1979;**44**:588–608.
148. Tolk A , Barros F, D’Ambrogio A *et al.* Hybrid simulation for cyber physical systems: a panel on where are we going regarding complexity, intelligence, and adaptability of CPS using simulation. In: *Proceedings of the Symposium on Modeling and Simulation of Complexity in Intelligent, Adaptive and Autonomous Systems*. Society for Computer Simulation International, 2018.
149. Kucukkaya G, Hester P. Maritime cyber security: system analysis and evolution of AIS. *Strat Cyber Def* 2017;**48**:160.