Educational and Information Technologies (2023) 28:11759–11791
https://doi.org/10.1007/s10639-022-11451-4

Cybercompetitions: A survey of competitions, tools, and systems to support cybersecurity education

Tyler Balon[1] **· Ibrahim (Abe) Baggili**[2]

Received: 1 September 2022 / Accepted: 2 November 2022 / Published online: 24 February 2023

Abstract

Over the last decade, industry and academia have worked towards raising students' interests in cybersecurity through game-like competitions to fill a shortfall of cybersecurity professionals. Rising interest in video games in combination with gamification techniques make learning fun, easy, and addictive. It is crucial that cybersecurity curricula enhance and expose cybersecurity education to a diversified student body to meet workforce demands. Gamification through cybercompetitions is one method to achieve that. With a vast list of options for competition type, focus areas, learning outcomes, and participant experience levels we need to systematize knowledge of attributes that ameliorate cybercompetitions. In the wake of the COVID-19 pandemic and global lock-downs, competition hosts scrambled to move platforms from local to online infrastructure due to poor interoperability between competition software. We derive a list of takeaways including the lack of interoperability between state-of-the-art competition systems, breaking the high knowledge barrier to participate, addressing competition type diversity, then suggest potential solutions and research questions moving forward. Our paper aims to systematize cybersecurity, access control, and programming competitions by surveying the history of these events. We explore the types of competitions that have been hosted and categorize them based on focus areas related to the InfoSEC Color Wheel. We then explore state-of-the-art technologies that enable these types of competitions, and finally, present our takeaways.

Keywords Cybersecurity · Hacking competition · Collegiate clubs · Gamification · Capture the flag · CCDC · CPTC

✉ Ibrahim (Abe) Baggili
ibaggili@lsu.edu

[1] University of New Haven, West Haven, CT, USA

[2] Baggili Truth (BiT) Lab, Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA

1 Introduction

Once a niche hobby, cybersecurity has grabbed the attention of government, academia, and criminals over the last 30 years. Cisco reported in 2018 that the average connections per household in North America was roughly 8.2 devices, this number is projected to grow to 13.4 by 2023 (Cisco Systems, 2018). The mass adoption of technology in home has introduced many new devices such as Internet of Things (IoT), which have resulted in countless security vulnerabilities and breaches (Abomhara & et al. 2015; Kolias et al., 2017). In 2008, The Center for Strategic and International Studies (CSIS) project was formed. CSIS sounded the alarm that the United States had a national security crisis involving cybersecurity following a wave of damaging cyberattacks (Langevin et al., 2008). The Obama Administration released a Presidential Commission in 2010 acknowledging the work by CSIS, citing that, "The cyber threat to the United States affected all aspects of society, business, and government, but there is neither a broad cadre of cyber experts nor an established cyber career field to build upon" (Evans & Reeder, 2010).

There was a call to academia to develop and enhance cybersecurity curricula. In 2013, the same administration introduced $62 million dollars in government funding dedicated to cybersecurity personnel and to expand the Scholarship for Service (SFS) program. This fund would provide an incentive for students to pursue a career in cybersecurity and become government employed post graduation (Office of the Press Secretary, 2020b). In subsequent years, there have been multiple CSIS briefings and related Executive Orders continuing the allocation of resources towards cybersecurity growth and development (Langevin et al., 2011; National Institute of Standards US Department of Commerce and Technology, 2020; Department of Homeland Security, 2020; Office of the Press Secretary, 2020a).

With the United States predicted to be approximately 62% short of cybersecurity professionals (Center for Cyber Safety and Education, 2019), there is still a large gap to fill with competent security professionals. Universities have put forward efforts to obtain new students that will study cybersecurity and achieve excellence in the domain. One promising methods which has been used to attract new students and expand educational outcomes has been to leverage the usage of Cybercompetitions (Eagle, 2013; Eagle & Clark, 2004). We define Cybercompetitons as: competitions or educational programs which teach computer science, information assurance, and network security through the usage of gamification and focused skill development in: programming, access control, and cybersecurity. We further discuss the different types of cybercompetitions, the components, educational outcomes, and present current shortcomings in Section 3.

The research community has examined the effectiveness of using events like Capture the Flag (CTF) to enhance cybersecurity education (Leune & Petrilli, 2017), and case studies which focus on positive learning outcomes from security training when gamification techniques are utilized (Thornton & Francia, 2014). Many competition hosts leverage gamification to better teach cybersecurity (Boopathi et al., 2015; Antonaci et al., 2017; Dabrowski et al., 2015). Higher education has employed gamification to improve information retention, willingness to study, and increased

self-learning (Schreuders & Butterfield, 2016). Younger interests have been attracted through creative themes like PicoCTF 2013 - Toaster Wars (Zhang et al., 2013).

The overarching purpose of this paper is to provide a valuable resource for future cybercompetitions to reference that will aid in creating the most effective educational outcomes for their participants. Our institution has immense experience in running cybersecurity boot-camps, and has participated in a large variety of competitions where we have had direct feedback from competition hosts and other competing teams. To the best of our knowledge, this is the first Systematization of Knowledge (SoK) paper related to cybercompetitions and our contributions are as follows:

1. Explore the history of cybercompetitions and build a taxonomy of the largely recognized competition types.
2. Provide an analysis of state-of-the-art systems that aid in the hosting of cybercompetitions by comparing and contrasting features, development plans, and the communities surrounding each technology.
3. Derive a list of takeaways and considerations for future researchers to continue to advance cybercompetitions and empower new individuals to host their own.

The rest of this paper is organized as follows. Section 2 gives the history of cybercompetitions, explores Cyber Clubs and competition hosts, briefly discusses gamification, and associates the InfoSEC color wheel to cybercompetitions. Section 3 presents a cybercompetition taxonomy, then goes on to: define the different types of competitions, skills for participants, and expected learning outcomes. Section 4 examines the state-of-the-art tools for building a competition and the shortcomings presented in the existing technologies. Section 5 provides a research narrative that includes a comprehensive review of the research currently in the domain, and we present takeaways and considerations to empower future research on cybercompetitions. Finally, Section 6 concludes the paper building a path for future work.

2 Overview

2.1 History of Cyber competitions

In the 1970s the first competitive programming events occured and the term ”hacker” was loosely applied to anyone with basic computing knowledge (Saman, 2007). In the 1980s academic conferences such as IEEE Symposium on Security and Privacy and USENIX Security Symposium began as a forum for computer security research. By the early 1990s cyber began to move past being only a niche hobby and academic area of study. Industry began searching for professionals with knowledge about information security and technology to help protect their companies from growing cyber-threats. Conferences began to help refine cybersecurity knowledge and educational efforts. Two conferences, DEFCON and BlackHat (Moss, 2008), would expand to host competitions alongside their security conference. These competitions kick-started a massive genre of events that would focus on information assurance, network security, and cybersecurity. In 1996, DEFCON hosted the first Cybersecurity CTF,

a first of its kind (DEFCON, 2018). Although this type of event grabbed the interest of many "hackers", according to the CTFWiki, the event was riddled with issues that caused frustrations for early contenders involving unclear rules for competitions, poor scoring systems, and badly designed platforms (CTFWiki Team, 2019).

Cybersecurity competitions didn't stop at CTFs, but extended into other forms of events. In the early 2000s, defense competitions became popular, which were designed to mock real world scenarios unlike CTFs. The Cyber Defense Exercise (CDX) for the five US service academies (Dodge & Ragsdale, 2004), and the first Collegiate Cyber Defense Competition (CCDC) from University of Texas in San Antonio (UTSA) (Conklin, 2006) were two of the first defense competitions created. Both competitions still exist to this day: CDX now being a part of the National Security Agency (NSA) and CCDC which has become a national multi-stage event (Fink et al., 2013). Defense competitions focused on teams defending a network or system, instead of attacking individual problems like that in CTFs. Until 2010, cybersecurity competitions were only designed and geared towards industry professionals and students in academia soon to enter the workforce, not to attract interests.

In 2010 President Barack Obama released a Presidential Commission acknowledging our nation was severely under-prepared to defend our systems from growing cyber-threats and called for academia to focus on teaching students to become cybersecurity professionals (Evans & Reeder, 2010). Following this Presidential Commission, Carnegie Mellon University (CMU) developed one of the largest competitions targeted at the high-school level, picoCTF (Chapman et al., 2014), with the goal of exposing cybersecurity to people at a younger ages. PicoCTF had major early success, with participation by tens of thousands of high-school students wanting to study cybersecurity (Chapman & Brumley, 2013). PicoCTF released their platform, scoring engine, and CMU released multiple publications outlining their experience. This encouraged other Universities to host CTFs, by utilizing the picoCTF platform, or by developing their own systems. Despite the issues showcased in the early DEFCON CTFs, the CTF genre has massively improved over the years, and today sees hundreds of variations hosted by numerous sources as seen on CTFTime[1], a platform for archiving CTF events.

2.2 Gamification theory

Gamification is a relatively new term, with research suggesting it was coined in 2008 (Deterding et al., 2011). A rudimentary definition states gamification is: *the use of game design elements in non-game contexts* (Deterding et al., 2011). Future work expanded this definition to note the mechanisms and goals of employing gamification (Scholefield & Shepherd, 2019; Schreuders & Butterfield, 2016). However, research suggested there has yet to be a standard definition (Seaborn & Fels, 2015), so for the purposes of this paper we will use the rudimentary explanation. Only as of the last decade has there been published research regarding the application of game design elements and its results applied to Computer Science and Cybersecurity.

[1]CTFTime (url: https://ctftime.org/)

There are several common game mechanics used as design elements that support user engagement (Zichermann & Cunningham, 2011). Further works refined their purpose and discussed the motives behind each element (Thornton & Francia, 2014; Blohm & Leimeister, 2013). Game design elements can be analyzed by the "MDA" framework – *Mechanics* (functions that make up the game), *Dynamics* (regarding the players' interactions), and *Aesthetics* (visual appeal and how the player feels) (Zichermann & Cunningham, 2011). We will briefly highlight the elements and their motives as follows:

- **Point Systems:** numerical unit indicating progress and allows participants to be ranked against each other.
  - *Motives*: self-achievement, social-recognition
- **Levels:** allows the participant to track the status of difficulty and progress.
  - *Motives*: self-determination
- **Leaderboards:** provides a ranking to players for comparison (*usually based on the Point Systems).*
  - *Motives*: self-achievement, social-recognition
- **Badges:** a token of achievement for completed tasks.
  - *Motives*: self-determination
- **Tutorials:** instruction to assist new players with on-boarding, facilitates capturing the participants attention.
  - *Motives*: intellectual curiosity
- **Challenges & Quests:** assignments and tasks for the participant to complete.
  - *Motives*: self-achievement, cognitive stimulation
- **Engagement Loops:** social interactions with other participants or hosts.
  - *Motives*: social exchange

Comprehensive research has reported that gamification produced positive effects and benefits (Hamari et al., 2014). Specifically applied to Computer Science and Cybersecurity, we have seen similar results when examining participant engagement, experience, and learning outcomes (Schreuders & Butterfield, 2016; Dabrowski et al., 2015; Scholefield & Shepherd, 2019). Many of the design elements which we have highlighted are incorporated into the state-of-the-art systems discussed in Section 4. One of the most notable cybercompetitions that have attempted to employ these elements is picoCTF in their 2013 competition: Toaster Wars (Zhang et al., 2013). The motives related to each element are what hosts attempt to leverage in cybercompetitions to attract new participants and retain their attention, while providing a new flair to education.

2.3 Cyber clubs, participants, & hosts

Similar to sports and other group orientated activities, teams have formed which solely focus on cybercompetitions. Events are composed of actors (can be individuals or teams) which engage an environment to solve problems to the host's specification (Sommestad & Hallberg, 2012). Hosts are responsible of deciding what type of competition to host, what types of problems to have actors solve, what the learning outcomes of the competition will be, how it is scored, and what rules are in place (Patriciu & Furtuna, 2009). Teams sizes vary per event and competitors participate with a wide variety of educational backgrounds. CDX consists of teams with 8 participants, whereas DEFCON has reported teams of 100 working together to solve the CTF challenges (Eagle, 2013). Although some events have team size constraints in place by the hosts, the competing team usually has numerous additional members, sometimes referred to as Cyber Clubs. Given the call from the United States government to expand focus on cybersecurity amongst academia and the opportunities for funding to Universities: many have created Cyber Clubs to increase educational outcomes and prepare a team of individuals to compete in these competitions for University exposure and student success (Cheung et al., 2012).

Cyber Clubs are usually student driven at the University level, but can consist of experienced faculty and staff that serve as coaches. Similar to sport teams: captains or leaders of the clubs will select their starting members, their backups, and have a group of participants that support in the venture and practice to prepare themselves to participate in the high level competitions. CTFTime is a platform that is used to not only advertise and archive events, but also rank teams on their success in these events (CTFTime, 2020). Cyber Clubs have become one of the biggest participants in creating CTFs as well as participating in them. Hosting a cybercompetition gives the club members experience in creating challenges as well as provides exposure for the University to other teams and schools. While Cyber Clubs do host a majority of the events, industry and government hosted events do take place, such as CDX (Dodge & Ragsdale, 2004), but usually industry and government will partner with events as sponsors to recruit new potential employees and encourage growth in the cybersecurity domain.

Some of the larger clubs have dedicated sub teams for the events which occur yearly, such as a Blue Team for CCDC (Conklin, 2006), a Red Team for Collegiate Penetration Testing Competition (CPTC) (Munaiah et al., 2019), and CTF teams (sometimes referred to as a Purple Team) for competitions like PicoCTF (Chapman et al., 2014), and CSAW (Cybersecurity Games & Conference) (Chung & Cohen, 2014; Gavas et al., 2012). This is seen in larger Cyber Clubs, as they have enough members to dedicate themselves to a different "color" of security (these colors are used to represent different cybersecurity disciplines, explained as The InfoSEC Color Wheel in Fig. 1).

2.4 The InfoSEC color wheel

There have been multiple iterations of a conceptual model known as the InfoSEC color wheel, which is used to explain security focus areas based on the primary (Red,

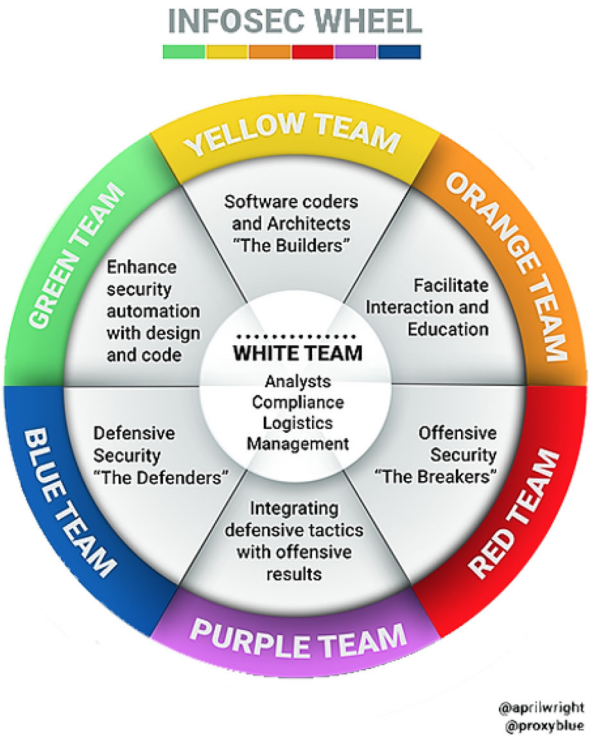

Fig. 1 Conceptual model of InfoSEC Color Wheel[4]

Yellow, Blue) and secondary (Purple, Green, Orange) colors. An entry that appeared in BlackHat USA 2017 identified the existing Red, Yellow, and Blue teams related to cybersecurity and introduced the terminology and expansion of Purple, Green, and Orange teams (April, 2017). This conceptual model, as seen in Fig. 1, is composed of seven colors which identify the different areas for security professionals. There are many different versions of this graphic, sometimes including more colors to identify the different areas, or laying them out in a different way, but these are the primary areas which are widely recognized in the domain. As the wheel has expanded, there have been mixed opinions from the security community as seen on Twitter and Reddit[2], that some of the entries are un-necessary and too specific.

The seven colors of the InfoSEC color wheel cover the different areas of cybersecurity in various ways (April, 2017; Cremen, 2020; Miessler, 2020). Each entry in the color wheel usually has competitions and problems designed specific to their knowledge, skills, and abilities. Those types of competitions are further explained in Section 3 and event types are highlighted in Fig. 2:

- **Red Team:** Known as the Attacker, focuses on Offensive Security and ethical hacking (Cremen, 2020). Consists of internal or external actors to a network,

[2]Reddit on InfoSEC Color Wheel (url: https://tinyurl.com/ybxrted2l)

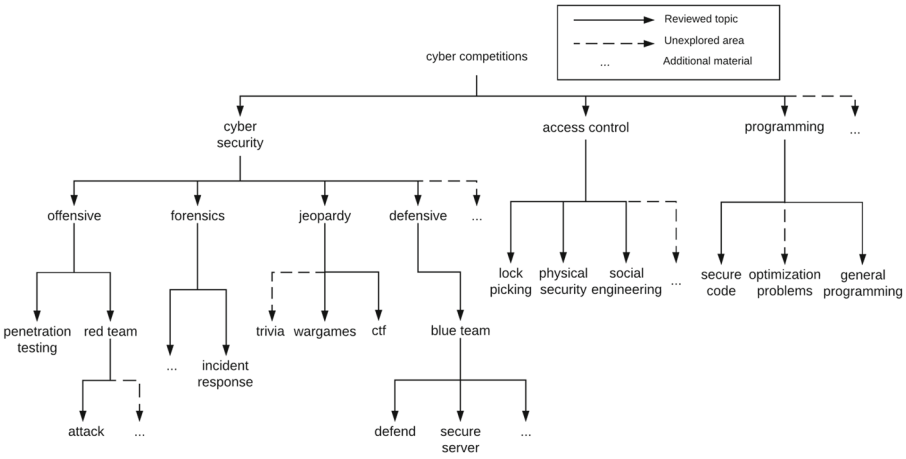

Fig. 2 Types of cybercompetitions broken into categories

which usually engages in controlled penetration testing (including white-box, black-box, and clear-box testing) (Brotherston & Berlin, 2017; Miessler, 2020). Red team members focus on exploiting vulnerabilities in systems and discovering new ones (known as a 0-day). Participates in Offensive Security events, Attack/Defend events, and jeopardy events. Competes against the Blue Team in events to target their network(s).

- **Blue Team:** Known as the Defender, focuses on Defensive Security and Incident Response (Cremen, 2020). Consists of internal actors on a network which work with intrusion detection methods, firewalls, and digital forensics to find actors attacking a network (Brotherston & Berlin, 2017). Blue Team members focus on protecting a network and doing damage control. Participates in Defensive Security events, Forensic events, and jeopardy events. Defends their network against the Red Team in events.
- **Yellow Team:** Known as the Builder, focuses on Programming and creation of code and systems (April, 2017). Consists of developers, software engineers, and architects which build code (April, 2017; Miessler, 2020). Participates in Programming events including general programming competitions, optimization problems, and secure coding. Works to build environments specified by administration.
- **Orange Team:** A cross between Red Team and Yellow Team, composed of individuals that help educate the Yellow Team on why the security vulnerabilities Red Team finds exist, and how to fix them April (2017). Orange Team members focus on secure coding, explaining why bugs exist, and how to improve insecure systems. Utilized as a team member for communication and education between Red and Yellow teams (Miessler, 2020).
- **Green Team:** A cross between Blue Team and Yellow Team, composed of individuals that help enhance blue team security tactics through code. Similar to the Orange Team, the Green Team assists in educating and assisting Yellow Team

members in implementing security features such as logging, automation, and monitoring that the Blue Team requires (April, 2017).

- **Purple Team:** A cross between Red Team and Blue Team, known as "hackers" which understand both Offensive and Defensive tactics in a system (Brotherston & Berlin, 2017; April, 2017). By understanding defensive tactics, they are able to better penetrate systems and vice versa. Participates in all of the event(s) that Red Team and Blue Team members participate in, as well as CTFs.

In addition to the teams listed, Fig. 1 mentions a 7th team known as the White Team which typically handles administrative tasks such as compliance, logistics, and management (Cremen, 2020). In a competition, the host is usually acting as the White Team. Different competitions may have an additional variety of colors associated to the responsibilities of participants and volunteers. CCDC, for example, utilizes the White Team as officials to observe teams participating, introduces a Gold Team responsible for organizing the entire competition, and a Black Team responsible for technical support during the competition (Williams, 2020). CPTC employs a White Team for competition management and a Black Team for technical support in the competition (Munaiah et al., 2019). The InfoSEC color wheel is to be used as an optional conceptual model. The model does not restrict nor define hard limits on the types of cybercompetitions.

3 Cybercompetitions

As cybercompetitions have grown over the last decade, the contents covered in each competition has expanded. More subjects in the cybersecurity domain have moved towards using competitions as a method to improve learning outcomes and provide a new flair to education. Competitions have shown promise in teaching participants how to work as a team and work in an intense atmosphere (Cheung et al., 2012), as some research has dubbed it "competing under fire" (Eagle & Clark, 2004). There are many different types of cybercompetitions hosted each year. New types of competitions are constantly being developed, some which have seen major success and become reoccurring events (like Rochester Institute of Technology (RIT)'s CPTC[5], one of the first to focus on penetration testing). There have also been Universities which entirely replaced a traditional class with competition based learning to study educational benefits and improved learning outcomes (Dabrowski et al., 2015).

In Fig. 2, the diagram breaks down different types of cybercompetitions and categorizes them into a tree. We categorized existing cybercompetitions into three major foci: Cybersecurity, Access Control, and Programming. Although a specific type of competition may fall into one sub-tree, they may have elements or aspects of that competition in another sub-tree (for example: a penetration testing competition in the cybersecurity sub-tree, might have physical security elements from the access control sub-tree). As explained by the key, the solid lines are topics and subject

[5]CPTC Competition (url: https://www.nationalcptc.org/)

materials which we focus on in this paper. The dotted lines and ellipses reference unexplored areas (implying that the tree is intentionally left incomplete, as there are other possible types of competitions and topics which can be used).

3.1 Cybersecurity competitions

Cybercompetitions focused on cybersecurity have been widely publicized, with CTFTime in 2018 and 2019 reporting well over 100 competitions per year advertised on their platform (CTFTime, 2020). Cybersecurity competitions have been documented and studied in academic research over the last decade, with numerous publications on building competitions, reviewing the learning outcomes, and surveying participants. Past work explored CTFs (such as PicoCTF, DEFCON, and CSAW) (Zhang et al., 2013; Cowan et al., 2003; Bashir et al., 2015) and cyber defense competitions (such as CDX, CPTC, and CyberPatriot) (Dodge & Ragsdale, 2004; White et al., 2010), many designed for the collegiate level of play (Chu et al., 2007).

Due to the complexity, Table 1 breaks down cybersecurity competitions and is laid out as follows: the *Description* outlines the topics that are covered and flow in each competition, and the *Learning Outcomes* explains the expected benefits from each competition.

Of the cybersecurity competitions listed in Table 1, CTFs remain the most popular, due to their ease of creation, hosting, and scaling. In general, cybercompetitions provide participants with skill development of tools, systems, and technologies. As explained in Section 2, the InfoSEC Color Wheel is a common theme in these competitions as there are specific events designed for Blue Team (ex: cyber defense, Attack+Defend, incident response) and Red Team (ex: penetration testing, Attack+Defend).

3.2 Access control competitions

Cybercompetitions have expanded to testing and teaching participants with not only cybersecurity but also physical security. Participants will be tasked to showcase their abilities in overriding access control mechanisms. These access control competitions have been hosted as standalone events (University of Connecticut, 2020), or in unison with events like CTFs (Cowan et al., 2003). Typically these events test how much access a participant can get (such as entering a restricted area), how many mechanisms they can defeat (such as breaking into an area), or what information they can obtain (such as social engineering). This is designed to replicate real world access control mechanisms. In Fig. 2, we describe the following types of access control competitions:

- **Physical Security:** events are designed to test participants knowledge of access control systems. The goal of physical security challenges are to gain access to a room or secure environment (such as a server cabinet designed to secure physical machine access). These events evaluate the access a participant is able to gain, the time it takes them to gain access, and if the participant is detected in gaining access. In some unique events, such as ToorCon, participants are asked

Table 1 Cybercompetitions related to cybersecurity topics

Competition Type	Description	Learning Outcomes
Offensive	-	-
Penetration Testing	Event includes a vulnerable environment for participants to attempt to compromise. Usually features software with publicly available Common Vulnerabilities and Exposures (CVE)ss, mis-configured systems, or other attributes found in vulnerable networks (Munaiah et al., 2019).	Ability to use penetration testing tools commonly found in Kali Linux. Experience in writing a comprehensive penetration testing report intended for both technical and non-technical staff and experience communicating and working with a third-party company.
Red Team *(attack)*	A King of The Hill (KOTH) type event (Eagle, 2013). Acting as an attacker, they will attack other teams using offensive security techniques (usually while simultaneously defending their own environment).	Offensive Security skills and Red Teaming skills, team management and coordination, utilization of CVEs, handling injects from a business or supervisor, delegation of tasks.
Forensics	-	-
Forensics	Event usually features a mock case or situation where evidence has been retrieved and must be examined using skills like Memory Forensics, File Carving, evidence analysis, etc.	Ability to examine evidence in a forensically sound manner. Experience using popular forensic tools like Autopsy, FTK Imager, Volatility, and SANS SIFT. Techniques to write a report and present data.
Incident Response	Event where host will provide logs, system images, and network traffic which a team will analyze then provide an analysis how to resolve issues in a compromised network. Events typically include a business response.	Knowledge of incident response workflow. How to prepare a report and assess cost of attack, downtime, and create recovery plans. Experience performing log management, traffic analysis, and using forensic tools.
Jeopardy	-	-
Capture the Flag (CTF)	Series of security challenges on the topics of, but not limited to: binary exploitation, reverse engineering, network analysis, forensics, web, and cryptography (Boopathi et al., 2015). Requires the participant to hunt for the ”flag” in the challenge (hidden in many different ways) (Eagle, 2013).	Each type of CTF problem requires the participant to use a different set of skills, thus teaching them a large variety of cybersecurity concepts. In addition, basic computing skills like usage of command-line tools, git, and Linux environments are learned.
Wargames	Similar to CTFs, more intended for beginners, a series of challenges which build on each other from basic cybersecurity concepts into more difficult challenges. Typically permanent websites that Cyber Clubs use to teach novice members.	Basic computing skills such as using the command-line, git, and Kali Linux. Typically features a variety of challenges but will teach from the most basic of problems into more complex ones (ex: basic binary exploitation techniques into more advanced ones).

Table 1 (continued)

Competition Type	Description	Learning Outcomes
Defensive	-	-
Cyber Defense	Simulation event where a designated (usually professional) Red Team will attack an environment given to participants (Fink et al., 2013). On these networks, the defenders must protect vulnerable software, resolve compromises, and correct mis-configured services. Objective is to emulate being a corporate network, typically.	Ability to detect threats against a network using log analysis and traffic analysis. Experience recovering from a compromise, re-configuring services, defending against an attacker, and removing persistent threats on a system. Experience in maintaining a large network of systems (usually in different sub-nets) and firewalls.
Blue Team *(defense)*	A KOTH type event (Eagle, 2013). Acting as a defender, participants will defend an environment from attacking teams (typically simultaneously attacking other teams). These events sometimes have a dedicated Red Team.	Defensive Security and Blue Team skills are gained, as well as team management and coordination, utilization of log analysis, network traffic analysis, real-time threat detection, and configuration of firewalls.
Blue Team *(secure server)*	Unique event where a team will prepare a server with secure systems, a hardened operating system, and any competition requirements (Cowan et al., 2003). The server will be put on a network for other teams to attack.	Operating system hardening techniques. How to configure software properly and defend against an experienced Red Team. Automatic threat mitigation, how to secure arbitrary services provided by a third party.

to attempt to break mechanisms in tamper-proof systems, as a means for testing industry standard equipment (Conti et al., 2011). In RIT's Information Security Talent Search (ISTS)[6]), participants are challenged with breaking into a secure server room. Past CPTC events featured physical infiltration techniques used to attack a computer (using a Universal Serial Bus (USB) and other devices). Physical security challenges typically incorporate Lock Picking to help gain access or to abuse a weakness in technologies such as Radio Frequency Identification (RFID) and magnetic strip authentication methods. Physical security events are designed to educate participants of potential weaknesses in thought-to-be secured environments.

- **Lock Picking:** is incorporated into competitions in combination with other challenges that all test the participants' ability to bypass locking mechanisms (Conti et al., 2011). There have been challenges in events which test the ability for a participant to break past a lock on a door to access a certain area, break free of handcuffs, or open a padlock. In ISTS, breaking out of handcuffs yields the participant and their team bonus points. In DEFCON, teams have been challenged to break into a box without damaging it or the locks (leaving minimal evidence

[6]ISTS Competition (url: https://www.ists.io/)

of tampering) (Cowan et al., 2003). Some events like MITCTF[7] have entire scoring categories based on lock picking problems. Lock picking is commonly used to test a participants' ability to gain physical access to an environment, but has been used as standalone challenges. Lock picking is used to educate participants mechanisms that prevent locks from being bypassed.

- **Social Engineering:** Social Engineering competitions test the participants' ability to divulge confidential information using deceptive techniques from a company or third party (Conti et al., 2011). Participants have been known to attempt to social engineer the competition host(s) for information related to flags or to get a competitive edge against other participants, although this can result in penalties against a team for breaking competition integrity. Attack vectors for social engineering include phishing emails, in-person interactions, phone calls, and leveraging live chat with company employees. In CyberSeed's Social Engineering Competition, the host used CTF-like challenges (asking participants to get certain information from employees) in a mock company setting, that was trained to release information when tactics were used (thus yielding the "flag" for the participant) (University of Connecticut, 2020). In DEFCON's Social Engineering Competition, participants research the target prior to the start of the event, then use reconnaissance techniques to collect information and leverage the target using social engineering (Conti et al., 2011). Social engineering events provide participants with experience in dealing with malicious actors who want to abuse psychology and target non-technical attributes to gain unauthorized access. There have been recent efforts to host a "pure Collegiate Social Engineering CTF" by Temple University (Temple University, 2020), similar to the one hosted by CyberSeed.

3.3 Programming competitions

Programming events and competitions have become more popular as there has been a shift for youth to "learn to code" (Kafai & Burke, 2014). There has been pushback acknowledging that not everyone may be able to code (Shein, 2014). However, websites, forums, internet communities, and discussion boards have made resources easier to access. Due to the learning curve of programming, many events keep their websites live post-event to be publicly accessible as a permanent resource. Programming events and competitions focus on general programming, secure coding, optimization, and mathematical problems. Unlike other competitions we've studied, programming competitions primarily focus on Computer Science theory, however, they can include some cybersecurity aspects:

- **General Programming:** challenges are designed for complete beginners to experienced programmers. General Programming problems test the participants' knowledge of basic program flow, logic, and concepts (Van der Vegt, 2006; Resnick, 2013).

[7]MITCTF (url:https://ctftime.org/ctf/180)

- **Secure Coding:** challenges a participant of both their programming and cybersecurity knowledge (Xie et al., 2015). The submission should address the challenge while ensuring code security. Submissions should not be able to be penetrated, abused, or otherwise compromised.
- **Optimization Problems:** focus on taking already written code or using advanced programming logic to speed up computational time and attempt to implement algorithms that out-perform other teams.
- **Mathematical Problems:** are similar to optimization problems, but focus slowly on taxing mathematical problems that are hard to compute.

Unlike other types of Cybercompetitions, not many resources exist to help hosts score participants. Most programming challenges typically use collaborative programming systems or version control to receive answers from participants and grade them. Collaborative systems, such as repl.it[8] are browser based IDEs which allow submitted code to be graded against a set of defined outputs. Version control such as GitHub or GitLab allow auto-compilation and auto-building of committed code, enabling the hosts to score against a local system. There have been proposed systems such as Code Hunt (Fouché & Mangle, 2015) which are specifically designed for scoring programming competitions, but have not been widely adopted. We further explore scoring engines in Section 4.

4 State of the art

There has been support by academia and industry to create state-of-the-art tools that assist in creating cybercompetitions, automate traffic generation, create advanced threats, and train the next generation of cybersecurity professionals. In this section, we discuss the tools which are used to aid in hosting cybersecurity competitions, examine the advanced technologies which assist in challenge creation, and explore shortcomings in the existing systems. Some projects have gone untouched due to lack of funding or interest. We will consider a tool actively maintained if it has been updated (new commits on git, new release, etc) in the last year.

4.1 Capture the flag (CTF) systems

A critical component to hosting a CTF is the scoring engine (also referred to as the scoreboard). Participants must have a platform where they can submit flags they find. Different types of software (platforms) have been released over the last several years which enables hosts to create a CTF without the need to develop their own platform. Early CTF platforms simply held scores for teams and accepted strings as flags, but as platforms have advanced they now accept coding flags, help with configuration, and add user management features (Chung & Cohen, 2014). A Facebook Research Engineer, apsdehal, created a repository of the most commonly used platforms and

[8]repl.it https://repl.it/

resources for creating a CTF known as awesome-ctf, which is actively maintained on github (apsdehal, 2020).

Table 2 lists CTF platforms (and their maintainers), used by some of the most popular CTF events according to CTFTime (such as picoCTF and CSAW).

According to a brief analysis of the CTFTime archive, the most commonly used CTF platforms have been FBCTF, picoCTF, and CTFd which we will review and analyze in depth.

FBCTF was created internally by Facebook's threat infrastructure team to host a college-level Facebook CTF starting in 2013. In 2016, the team at Facebook decided to make the software open source, noting: "due to a high cost and technical requirement of building and running a CTF, [and] few publicly available resources" (Singh, 2020). FBCTF is available on github, but is no longer maintained, notably having 100+ open issues which have gone unsolved (Facebook, 2018). The software used to be included in Facebook's Bug Bounty Program, but has since been removed. The repository is currently archived, with no indication of future maintenance or development. The platform is written in PHP and suffers from scaling issues (Lei et al., 2014). FBCTF is still a contender for smaller CTFs due to its' ease of setup and visually appealing world-map and "hacker" like theming.

picoCTF was created by students at CMU in 2013 for their own competition but was released publicly. The platform has gone through several revisions by students and by company sponsored changes to add features and security enhancements to the platform. As of 2020, picoCTF is available on github, and is currently maintained (Carnegie Mellon University, 2020a). The platform is written in python and has the following notable components:

- *Web server (and API):* hosts the problem list (including hints for problems), scoreboard, classroom feature (to make the platform usable in a classroom setting), and the shell server integration (allows an in-browser shell experience).
- *Shell manager:* lowers the overhead of CTF configuration by a central point to manage server(s) needed for the competition, and aids in providing competitors access.

Table 2 Capture The Flag (CTF) Platforms

CTF Software	Maintainer	Reference
FBCTF*	Facebook	(Facebook, 2018)
picoCTF*	CMU	(Carnegie Mellon University, 2020a)
CTFd*	Kevin Chung / NYU	(Chung, 2020a)
MITRE CTF	MITRE	(MITRE, 2020a)
echoCTF	Echothrust Solutions	(Echothrust Solutions, 2020)
LibreCTF	EasyCTF	(EasyCTF, 2019)
NightShade	UnrealAkama	(UnrealAkama (avidhacker), 2017)

*denotes most popular software

- *Problem manager:* picoCTF's efforts at standardizing creation and deployment of CTF challenges so they can be re-used.
- *Ansible deployment:* allows for installations and configuration to be automatically deployed through the popular DevOps tool, Ansible (Hochstein & Moser, 2017).

The platform has gone through 12 official releases and thousands of commits. PicoCTF is feature-rich and has a large list of contributors. Compared to CTFd, picoCTF falls short in marketing, naming itself after it's notable competition and not a standalone platform, likely contributing to why it is not as popular as CTFd.

CTFd is developed and maintained on github by Kevin Chung, a former New York University (NYU) Graduate Student as a result of the popular NYU CTF: CSAW (Chung, 2020a). CSAW is a "cybersecurity games and conference" venue hosted by NYU annually. CTFd's creator[9], expressed that there was a lack of a powerful CTF platform and pitfalls in CTFs themselves, such as high entry level, that CTFd aimed to help solve (Chung & Cohen, 2014). CTFd started as a small CTF platform that only supported a scoreboard to now, an extensive feature-rich "framework":

- *Multiple Challenge Types:* unlike other platforms, CTFd supports advanced challenges opposed to the traditional 'submit a flag' structure.
  - **Standard** - traditional flag submission
  - **Coding** - integrated programming similar to repl.it
  - **Multiple Choice** - useful for quizzes or trivia
  - **Manual Verification** - host flag approval
  - **Decaying Points** - lower points over time
- *Advanced Scoreboard & Statistics:* provides a robust experience by breaking down solves by individuals, teams, providing tie-breaking features, allowing scoreboard freezing, and players to spend points for hints.
- *Team Management:* allows users to control their own teams, compete as individuals, join and leave teams.
- *Page Manager:* enables the host to use CTFd as an entire website with custom pages for information, legal notices, and other uses.
- *Plugin Manager:* allows the community of users who use CTFd to run their competitions to actively contribute by writing their own plugins (Chung, 2020b). Some plugins have made it into the master code-base as a feature of the platform. Some of the most notable plugins featured include:
  - **Portable Challenges** - enables rapid deployment of challenges in CTFd and editing outside of the platform (Texas, 2020)
  - **CTFd Docker** - allows hosts to create, edit, and manage docker containers from the CTFd interface seamlessly (Kevin, 2018)

[9]CTFd About (url: https://ctfd.io/about/)

- **CTFd Hash Crack King** - extends the types of challenges, adding a King of the Hill (KoTH) plugin to score teams controlling a server (Tyler, 2018)
- **Online Challenges** - allows dynamic flags and cheating reports to aid in finding teams sharing flags and ensure competition integrity (XuCcc, 2018)

CTFd is written in python and is scalable. It has taken over the event list on CTFTime, being the most commonly used platform for CTFs. This is a result of the community around CTFd and the platform's marketing. The plugin manager sets CTFd apart from other platforms mentioned as it enables community members to write their own features and share them without the need to wait for a merge on GitHub. CTFd, started as a platform for CTFs and has gone on to be a business not only releasing their own software but providing hosting, custom themes, custom plugins, and competition management to create CTFs "as a service".

Although there is a long list of contenders, no platform is perfect. One of the long lasting issues in CTFs is there is no formal format for creating challenges, therefore moving from one platform to another is a tedious task and requires the host to essentially start from scratch. There have been attempts at solving this, like picoCTF adding a problem manager, but until the majority of platforms accept a standard format, this will not be solved.

In addition to issues with platform interoperability, there are also security concerns which have appeared over the years. During RC3 CTF 2016, a user known as "seadog007" used a vulnerability in CTFd to mass generate points and take over the scoreboard (Yu, 2016). Also in 2016, during CODEGATE CTF a user leveraged privilege escalation (CVE-2016-1576) to gain root access on the system and view all flags intended to be found via challenges. While it is usually in the rules of cybercompetitons not to attack competition infrastructure, it does not absolve these platforms from responsibly securing themselves against potential adversaries. If a competition is hacked, it can damage the host's credibility bringing into question their ability to teach cybersecurity topics. NIZKCTF aimed to solve this by defining a competition protocol and protecting flags, but the platform was not widely accepted (Matias et al., 2018).

4.2 Competitions "At-Scale"

For competitions that do not follow the CTF format, such as the Offensive and Defensive entries in Table 1: advanced server infrastructure and scoring engine capabilities are required to run these types of competitions. We refer to these competitions as "at-scale" due to their size and scalability requirements for the server infrastructure. Competitions achieve this advanced server infrastructure by leveraging virtualization software, configuration management, capitalizing on new technologies such as cyber ranges, and working with "the cloud". Though, each solution is challenged with different issues. In this subsection, we examine the leading solutions to achieve competitions "at-scale" by exploring their usage capabilities, their cost, and the various issues, challenges, and concerns that come with each technology. We will see that

in practice, these solutions are usually used in combination to create competitions "at-scale".

The cloud has revolutionized the Information Technology (IT) industry (Dillon et al., 2010), moving locally installed software and systems to the internet users can connect from an essentially unlimited number of locations (Hayes, 2008). National Institute of Standards and Technology (NIST) has defined cloud computing as, "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort" (Peter et al., 2011). In general, cloud computing can be cost effective with services like DigitalOcean[10], Azure[11], and Amazon Web Services (AWS)[12] starting at only a few dollars per instance, and offering significant price reductions and account credits to University, Institutions, and open source projects. This scale on-demand and ease of deployment solution has been beneficial to competition hosts who require scalable and remotely accessible infrastructure.

Although cloud computing is a notable contender in solutions for competitions "at-scale", as with any newer technology there are challenges presented as the technology is widely adopted. Several researchers have disclosed challenges with cloud computing (Dillon et al., 2010; Popović & Željko, 2010; Sajid & Raza, 2013), all which put competition integrity into question. Research has suggested that security, privacy, performance, reliability, and availability can all be compromised in a cloud computing environment. Further research has suggested there jurisdiction concerns when performing cloud forensics in response to these attacks (Ruan et al., 2011), further complicating the configuration. As these cybercompetitions have contestants who are highly skilled in ethical hacking, cybersecurity, and digital forensics – there is potential for cheating and the leveraging of vulnerabilities in cloud software. There are also concerns with costs related to cloud computing. While standalone cloud services are cost effective and smaller competitions can bear the burden, larger competitions can cost tens of thousands of dollars for an event lasting a few days. Researchers have developed decision-making models to aid in cost analysis of cloud computing and optimization methods (Chaisiri et al., 2011; Martens & Teuteberg, 2012), but as a competition scales the bill can grow exponentially. There have been several attempts at making private cloud solutions, such as Anstle[13], but these too have their shortcomings. Cloud computing is a notable contender, and in practice is used with other solutions we highlight in this subsection.

Virtualization is an abstraction of computing resources, with the purpose of improving resource utilization (Sahoo et al., 2010). It is used across the industry to allow companies to virtualize instances of software and leverage the improved resource utilization to make their physical systems more cost effective. Virtualization enables enhanced software interoperability and platform versatility between hard-

[10] DigitalOcean (url: https://digitalocean.com/)
[11] Azure (url: https://azure.microsoft.com/)
[12] AWS (url: https://aws.amazon.com/)
[13] Anstle (url: https://antsle.com/)

ware and software (Brooks et al., 2012), therefore it has become widely popular. "Type 1" hypervisors, such as Xen, Microsoft's Hyper-V, Proxmox, and VMWare sit between the installations and the hardware, virtualizing installations. "Type 2" hypervisors, such as Docker, run on an already installed Operating System (OS) and containerize the system. There are notable differences between Hosted, OS, and Bare Metal virtualization modes: but those are out of the scope of this paper. Competitions use virtualization for ease of re-deployment, cloning of Virtual Machines, and the ability to install a large array of OS on a single physical host. It enables a competition host to make a single environment and duplicate it as necessary for competitors. Backup functionality in virtualization software is extremely useful should a host need to revert an environment.

There are several shortcomings with virtualization when observing the different cybercompetition use cases. While it is useful in smaller in-house competitions, remote competitors require a Virtual Private Network (VPN) to interact with the environment as if they are on the internal network. This complicates the configuration and adds another layer of potential attacks. Competition hosts have used virtualization in the cloud, but there are security concerns with cloud virtualization. In recent years, there have been multiple disclosed security vulnerabilities affecting Intel and AMD processors which impact virtualization security. Spectre (Kocher et al., 2019) (CVE-2017-5753 & CVE-2017-5715), Meltdown (Lipp et al., 2018) (CVE-2017-5754), and Foreshadow (Weisse et al., 2018) (CVE-2018-3615) all impact processor security, allowed attackers to escape Virtual Machines and exfiltrate data. While there have been security patches for these vulnerabilities, it has reduced processor performance, especially in cloud environments.

Configuration Management (sometimes referred to as Infrastructure as Code (IaC)) is an emerging technology that the industry is adopting at a rapid pace. Configuration management enables systems to be automatically provisioned based on code instead of a manual installation process (Benson et al., 2016). While competitions may configure a single system, or set of systems, one time and clone it for deployment, this is not a favorable approach as it is exceedingly difficult to move systems across different hardware, cloud computing companies, and virtualization software. There is a growing list of software which achieves IaC, but the following are notable systems that have been used to configure environments for competitions:

- **Ansible** - *python/YAML*, easy to use
- **SaltStack** - *python/YAML*, easy to use
- **Chef** - *ruby/DSL*, more difficult to use
- **Puppet** - *ruby/DSL*, more difficult to use
- **Terraform** - *go/HCL*, extremely difficult to use

It is important to note that terraform, unlike the others listed, is more of an "orchestration tool" instead of configuration: which means it is designed to provision the server and then leave the configuration job to other tools. This sometimes results in chef or puppet being used in combination with terraform in practice. In this instance, terraform, would provision and install the server, but chef or puppet would configure it. There are other trade-offs to consider when selecting a configuration manager such

as if you want a master v. master-less, agent v. agent-less, and procedural v. declarative: but these options are out of the scope of this paper (Yevgeniy, 2016). One of the major drawbacks using IaC is the complexity. While Anisble and SaltStack are considered easier to use and more reasonable, they are still difficult for the everyday user or professor that is not highly technical. Terraform on the other hand goes as far as deploying its own configuration language, complicating things one step further. In order to package these technologies in an intuitive manner, cyber ranges have been created and deployed.

Cyber Ranges are interactive, simulated representations of an organization's network (including systems, tools, and applications) that are connected to a simulated internet designed to reflect a realistic situation (National Institute of Standards and Technology, 2020; Paulsen et al., 2012). Cyber ranges are used and provided by the government, academia, and industry resources. The National Cyber Range (Ferguson et al., 2014), an innovative from the Department of Defense, is one of the most notable cyber ranges in the country and has attracted the attention of major universities like John Hopkins University and US based corporations like Lockheed Martin to create a joint effort of developing a scale model of the internet. There has been a growing interest in cyber ranges as they provide a virtual environment to use for cybersecurity training and exercises. A cyber range gives a safe, legal environment to develop cybersecurity skills and has proven to be extremely useful for competition hosts. Research has been published in designing games that raise awareness in cybersecurity and help raise "adversarial and system thinking skills" (Kianpour et al., 2019).

The most significant drawback of a cyber range is the entry price point and the price to maintain the range year to year. The National Cyber Range has been a multi-million dollar venture of up to $93 million (US Department of Defense, 2019). One of the largest academic cyber ranges is hosted by Virginia Tech, comes in at about $4 million (Sabbath, 2020). For smaller Universities and hosts, these numbers would be extremely difficult to achieve. This leaves the options of purchasing from a company that specializes in creating cyber ranges, or by building ones' own. Cyberbit, a company located in Israel that specializes in selling cyber ranges, estimates that to build your own range it will cost between $336,000 to $576,000 with an average yearly upkeep of $85,000 (Cyberbit, 2018). Cyberbit, as with most companies that offer cyber ranges as a service do not list their prices and require a custom quote that increases depending on the number of students needing access.

4.3 Automation & challenge creation

One of the largest obstacles present in hosting cybercompetitions is the creation of problems can be expensive and taxing (Newby, 2018). In most cybercompetitions, participants are competing to out-rank other teams or obtain prizes. This provides limited availability for problem re-use, and re-use of generated traffic data. Ideally, to enhance educational outcomes for participants, challenges should reflect real-world scenarios. This burdens the competition host with the task of preparing unique problems, creating fake traffic, and generating mass amounts of data for each event.

To promote fairness for competitors, problems should be unique across events, and annual occurring events should have unique data year to year.

It has been identified that one of the largest barriers to hosting these events is cost. Costs are generally associated with, "(1) hardware costs for hosting the competition, (2) the human resource expense required to administer the competition, and (3) the availability of and/or investment associated with competition material for the particular event" (Taylor et al., 2017). Due to the high cost in content creation, some events have opted to deny participants from competing over multiple years (Ncac - faq, 2019). This is a poor solution as it prevents participants from learning and teams from practicing on previously hosted events. To help reduce the high costs, automation and advanced challenge creation techniques must be developed.

To date, there have been a variety of solutions developed to tackle issues associated to cost and scalability of cybercompetitions. We will first explore tools created by academia. There have been several solutions created to support "cyber simulations" and improve "cyber wargames" (Bumanglag et al., 2019), with the largest contributor to this pool being Carnegie Mellon University (CMU, the hosts of the widely popular PicoCTF) (Mayes, 2020):

- **TopGen** is an application-service simulator. The tool assists in generating an internet simulation by allowing a single host (physical machine, VM, etc) to serve multiple co-hosted virtual services (HTTP vHosts, DNS views, SMTP/IMAP virtual mail domains). The code repository is currently maintained and can be found on github (Carnegie Mellon University, 2020b).
- **GreyBox** is a single-host internet simulator. Greybox allows for a single host to simulate an "illusion of connectivity", by providing a realistic BGP backbone topology, realistic latency by point-to-point link delays based on physical router distances in real-world locations, and a variety of website types emulated. The code repository is currently maintained and can be found on github (Carnegie Mellon University, 2020c).
- **GHOSTS** is a framework for NPC traffic emulation. GHOSTS automates and orchestrates non-player character activities on a network to provide simulated traffic for cyber exercises. GHOSTS will realistically mimic behavior of different users found on office and enterprise networks by running commands, browsing websites, accessing systems, creating documents, and more. The code repository is currently maintained and can be found on github (Carnegie Mellon University, 2020d).
- **vTunnel** is software which removes administrative traffic from a cyber exercise network. It aids in competition hosting, by assisting the host's scoring engine, and any command servers which must send commands to competitor's machines. Traffic is tunneled from a guest VM through the hypervisor allowing administrative network activity to be hidden. The code repository is no longer maintained, but can be found on github (Carnegie Mellon University, 2018).
- **WELLE-D** creates virtual wireless network environments. WELLE-D emulates 802.11 wireless communications virtually to aid in learning about wireless network security. It creates and passes 802.11 frames but hides them from the wired connection to allow realism of wireless communication. Multiple VMs can be

connected wirelessly to each other (virtually). The code repository is no longer maintained, but can be found on github (Carnegie Mellon University, 2019).

- **TopoMojo** simplifies virtual lab creation and development by providing a simple web application to hosts. It is a Linux based virtual appliance that launches VM learning environments. The goal of TopoMojo is to simplify the setup process of cyber exercise labs, allowing for easy deployment and re-use. The code repository is currently maintained and can be found on github (Carnegie Mellon University, 2020e).

Academia is not the only contributors to automation techniques. Industry has contributed heavily with various tools built to simulate red team and blue team activity. MITRE maintains an active research project known as the ATT&CK framework (Strom et al., 2018), which focuses on network defense and adversaries. From this project, multiple applications have been constructed such as CALDERA (MITRE, 2020b), CASCADE (MITRE, 2018), and CAR (MITRE, 2020c). CALDERA focuses on allowing autonomous breach-and-simulation exercises to simulate red team engagements, while CASCADE assists the blue team in automating investigation work. CAR on the other hand is a knowledge base of analytics based on the ATT&CK framework. This suite of projects can be used in simulations and competitions.

4.4 External resources

Cybercompetitions reportedly have a high knowledge barrier that prevents students from participating (Cheung et al., 2012). Once a cybercompetition concludes, feedback provided to participants from hosts is minimal. Participants are required to discover external resources which prepare them for cybercompetitions, train them for certifications, and prepare them for the workforce. There are resources available at a learner's disposal:

- **HackTheBox** – online platform for penetration testing labs. Offers a free option and a VIP subscription with added features and resources for roughly $11.99/month.
- **TryHackMe** – community based platform for wide range of cybersecurity topics. Offers a free option with limited learning materials and a subscription for roughly $10.00/month.
- **VulnHub** – community based platform with 'hands-on' experience in cybersecurity topics. No subscription model. All self hosted materials.

In addition to online platforms dedicated to cybersecurity education, there are a plethora of websites that have self learning paths to teach programming, secure coding, forensics, and other topics outlined in Table 1. The community based external resources where users can submit their own materials for others to utilize appear to succeed far more than websites offering a very restrictive set of learning materials at a high price point.

5 Research & discussion

5.1 In literature

We examined a large array of papers dating back to 2000 when cybersecurity competitions were a novelty and many of the now well-defined competitions like picoCTF (Zhang et al., 2013) and CCDC (Conklin, 2006) were in their early iterations. The paper (Zhang et al., 2013) defines and designed the yearly highschool CTF hosted by Carnage Mellon University (CMU). In Chapman et al. (2014) and Chapman and Brumley (2013), and Owens et al. (2020) CMU analyzes the learning outcomes of the competition by breaking down the problems which students favored, and how year to year CMU works to improve student engagement and learning outcomes. A later generation of CMU students updated the list of publications based on picoCTF with Burket et al. (2015): a paper which focuses on defeating cheating in CTFs through developing methods for automatic problem generation. The authors disclose statistics and demographic information of picoCTF, then release their problem set and platform for other competition hosts to utilize. The picoCTF competition to date remains one of the most popular competitions and one many others base their design on.

Since 2010, there have been numerous new programs and additional support by Universities to instruct students on security related topics. In Cheung et al. (2012), the authors focus on the lack of experience and expertise students have in cybersecurity and suggest that competitions have a high knowledge barrier to participate in them. They collect metrics from their own students by hosting an internal competition to analyze when students participate and what deters them from participation. The authors find that most of the students participate late in evenings, and that classes and other University activities prevent full engagement. Authors in Schreuders and Butterfield (2016), Dabrowski et al. (2015), and Boopathi et al. (2015), and Yonemura et al. (2017) focus on designing their own in-house competitions as a supplement to replace typical classes utilizing gamification theory (design theory, point based grade structure, and quests/paths for achievements) to gauge if it improves student participation and learning outcomes. They find that gamification improves learning outcomes and retention of information learned. In Woszczynski and Green (2017), the authors focus on learning outcomes specifically for Cyber Defense competitions, such as CCDC, by surveying judges and mentors of these types of competitions. Based on the responses they build a set of recommended outcomes expected for cyber defense competitions based on the importance of topics and participant preparedness.

The authors in Tobey et al. (2014) attempted to host the National Cyber League (NCL), and felt that cybersecurity competitions should be like sporting events and host regular 'hack-a-thons'. The paper broke down the gender ratio of participants and focused on how if students struggle, they become more likely to withdraw from the competition. In Conklin (2006), the authors define the major national competition still held today, CCDC. The paper presents the methodology behind how the competition was ran the first few years of its' existence. The papers (Chu et al., 2007) and (Dodge & Ragsdale, 2004) focus on improving design tactics and organization

cyber defense competitions based on elements utilized by each host. In Dodge and Ragsdale (2004), authors provided information about the Military Academy CDX, and described the scenario used in the competition and what learning outcomes are desired for participants. The paper (Chu et al., 2007) discussed attack/defense competitions, reviewing how the complexity of the competitions grows as it scales participation. The authors review iCTF and CCDC from 2005 and 2006 based on the scale, content, complexity, scoring, and rules of the competitions.

Throughout existing research, there has been a major focus on attracting new participants and how to retain them. One solution which many competitions have deployed is to add more game elements, leveraging the psychological theory of gamification. This method has been seen across research covered in this paper, as more and more competitions over the years have moved to having game like elements (achievements, scoreboards, etc.) and even game themes (competitions have been themed off the popular TV Series Mr. Robot[14]). The papers (Seaborn & Fels, 2015; Thornton & Francia, 2014; Boopathi et al., 2015), and (Schreuders & Butterfield, 2016) all apply gamification methods to cybersecurity competitions then document the success of the participants and feedback they received. In Seaborn and Fels (2015), the authors provide a comprehensive survey on gamification theory and implementation, related to Computer Science and STEM majors. The paper explains the taxonomy of game design, what frameworks are successful, and reviews the works of Universities utilizing gamification to improve higher education. The authors in Thornton and Francia (2014) focus on using gamification theory to improve the learning outcomes from their Cybersecurity students. The authors use a case study that took place over multiple semesters and had respondent data of over 300 students, with the majority of students saying gamification improved their learning experiences and declared what aspects they favored. In papers (Boopathi et al., 2015) and (Schreuders & Butterfield, 2016) the authors build their own environments for students to learn Cybersecurity through gamification. They replace traditional grading scales in their classes with experience systems and achievements, and report the student's successes.

Overall, there has been a profuse attention on bringing gaming into the classroom to support students learning difficult topics like cybersecurity. Ultimately, academia has focused on developing new ways and new design criteria to improve participation and learning outcomes. According to literature, learning outcomes to have the attention of researchers as via gamification: students have been proven to learn more and prefer ditching the traditional classroom structures. This can be seen as beneficial both students and institutions, as gamification empowers students to put in extra time and effort to reach self-made goals and collect achievements. To date, there has been many papers in the domain which focus on competition design, gamification theory, and survey data from students which provide metrics on learning outcomes. However, there has not been a comprehensive paper systematizing the knowledge related cybersecurity competition. Our work aims to fill this gap and provide takeaways to the reader.

[14] https://www.usanetwork.com/mrrobot

5.2 Takeaways

Over the course of this research, we studied over 100 papers and resources related to cybersecurity competitions. We have amassed a list of takeaways and considerations for future research to help improve cybercompetitions and develop new technologies to aid in their hosting. It is imperative for future work to continue to attract new students to cybersecurity and help preparing our future workforce.

Key Takeaway: Insufficient usage of learning outcomes results in competitions lacking real world scenarios and skill development. While we provided an outline of *expected* learning outcomes associated to competition type in Table 1, most hosts define their own outcomes. This, in combination with the majority of competitions being CTFs, results in a lack of learning outcomes which are useful for our workforce. Cybercompetitions should reflect industry workforce skills. One potential solution that avoids redesigning the wheel, would be to incorporate already widely recognized frameworks such as the NIST National Initiative for Cybersecurity Education (NICE) Framework (Petersen & Santos, 2020). NICE is designed to provide a common set of building blocks which organizations can draw from. The original iteration featured 7 categories of study, 33 specialty areas, and 52 work roles. The framework has since been updated to reflect a final release as of November 2020.

Key Takeaway: Cybercompetitions have a high knowledge barrier for participants. This deters potential participants from developing skills in cybersecurity. Competitions such as CSAW have released resources such as CTF101[15] to help prepare students for competitions. Other competitions simply link materials that *"could"* help, however, this stuns preparation. The NICE Framework (Petersen & Santos, 2020), provides an extensive information on specialty areas and skills required to fill work roles. Ideally, competitions could utilize this to align the outcomes based on what skills are required to fill a work role. Optionally, competitions could run alongside educational lectures that would also help participants develop cybersecurity skills.

Key Takeaway: There is a lack of diversity between competition types. As we've outlined, the majority of competitions are CTFs. Secure coding, forensic, incident response, and penetration testing are under-numbered. Competition hosts need to be more open to sharing resources and publishing works based on their competitions. By sharing more resources on how competitions are designed, what problems have been designed, and effectively how to "do it yourself": this will solve the lack of variety. Our work has been written to collectively talk about outcomes, tools, and resources to help close this gap. There is room to develop new competition types which support multiple foci in Fig. 2. The Defense Advanced Research Projects Agency (DARPA) Cyber Grand Challenge (Song & Alves-Foss, 2015; 2016) implements autonomous cybersecurity research based on Offensive and Defensive attributes through Program-

[15] CTF101 (url: https://ctf101.org/)

ming and elements such as Artificial Intelligence (AI) and Machine Learning (ML). Teams build their own autonomous systems (Nguyen-Tuong et al., 2018) that attack each other, then participate in an automated CTF, if they qualify. New competition types, such as these, will require further research in scoring engines and applied gamification, however, they provide participants with a unique experience in a variety of disciplines.

Key Takeaway: State-of-the-art software has low interoperability. Interoperability affects the majority of the technologies discussed in Section 4, especially CTF software. Low interoperability creates complications for hosts that need to migrate technologies and scale on demand. This frequently limits the amount of students who can participate and benefit from a competition (as seen in CPTC). PicoCTF (Carnegie Mellon University, 2020a) has taken a step in the right direction with their *Problem Manager* that enables ease of flag re-use, but more CTF systems need a similar approach. Our lab annually hosts the GenCyber Agent Academy – a K-12 program to teach cybersecurity to young students (Ladabouche & LaFountain, 2016), and have had complications moving from FBCTF to CTFd when FBCTF poorly scaled to meet demand. It is a responsibility of the communities surrounding these CTF systems to build extensions to standardize flag format or to write conversion tools from other popular software. Similar import systems are seen in bulletin board/forum software to allow conversion from one software to another seamlessly.

Takeaway: Competitions should encourage remote play. Previous to the 2020 COVID-19 pandemic, competitions frequently had in-person portions. There was a lot of difficulty in adjusting to lock-downs and in-person gathering restrictions, resulting in events to be cancelled or delayed. As interoperability of moving systems increases (such as putting a local system to the cloud easier) and costs continue to reduce across the board: competitions should attempt to allow remote participants. Researchers have proposed EZSetup (Li et al., 2017), a tool that aids in creating and managing virtual cloud environments for cybersecurity practices to help make using the cloud easier. Besides public health crisis there are other reasons why a participant may not be able to be on-site for an event, such as seen in (Cheung et al., 2012) where students would avoid competitions due to high workload in their normal classes. For portions which may require physical attributes (such as entering a server room), competition hosts should explore as a potential solution as a potential solution (Hassenfeldt et al., 2020).

Takeaway: Scoring methods lack standardization. CTFTime has attempted to solve this this problem by providing team ratings based on points gained in each competition. Participants themselves can vote on the difficulty of the competition and provide feedback to the hosts. To assist competition hosts in assigning point values to problems, there should be a widely accepted scale to determine the worth of each type of problem. For example, a Reverse Engineering (RE) problem, where most Universities have RE courses as 400 level or higher, should be worth more than a web challenge, where most Universities have web courses as 200 level courses.

Takeaway: Automated scoring will aid competition integrity. For problems that cannot be analyzed using a point scale based on difficulty, such as final reports in

competitions, there is room for error as graders can: (1) miss key information, (2) have bias towards a specific team, (3) use subjectivity based on findings. Scoring could be automated by leveraging deep learning and language processing techniques. Hosts would maintain a database of potential findings in a competition and the system could score against it. Research has occured presenting a framework for evaluating the exploitability of vulnerabilities (Moscovich et al., 2020), which could further help assess the worth of findings. This would help reduce bias, providing a grader an unbias score to reference to ensure their grading is equally applied to every submission.

Takeaway: There is a lack of communication between hosts and competitors. Hosts will frequently deny releasing information about infrastructure or a competition to allow re-use in future years to avoid additional costs. As we continue to develop new state-of-the-art systems which help reduce the cost of hosting cybercompetitions, we would enable hosts to not need to re-use as many materials and open discussions about each competition. CPTC has moved in this direction by providing more feedback to teams, releasing the grader's notes, sharing team reports publicly (Morris, 2020), and by speaking at conferences on their competitions. This openness has led to higher learning outcomes, whereas feedback in earlier years was overwhelmingly that the competition did not share enough details to help students learn. Other competitions should take this approach in enabling skill development.

Takeaway: Configuration management and IaC are cryptic and difficult to maintain. When automation is rendered difficult to deploy, it deters hosts from utilizing the many benefits to configuration management and IaC. National competitions must rely on these solutions to deploy enough instances for participants to use during a competition, but smaller events may not be able to benefit from these systems. Hosts who have successfully used configuration management and IaC should be encouraged to publish their work, of not only how the competition was designed, but how these state-of-the-art tools have been employed to achieve success. IaC tools also can be developed to directly support cybercompetitions, laforge (Levinson, 2020b), a Security Competition Infrastructure Automation Framework, provides rapid development of infrastructure for cybercompetitions. More publications and projects such as these will empower more hosts to benefit from these technologies.

Takeaway: Cyber ranges are not widely adopted. The high cost for creating and maintaining cyber ranges deter them from being widely used for cybercompetitions. Besides initial costs to implement a cyber range, scenarios can cost thousands to create. Additional automated attack frameworks, such as MITRE ATT&CK will benefit cyber ranges by reducing the need to add adversary activity into scenarios manually. This can be achieved by leveraging ML and AI to deploy already publicly disclosed CVEs and automatically compromise systems. Work has been presented at DEFCON to showcase Deep Exploit (Levinson, 2019a), a tool for automated penetration testing using Deep Reinforcement Learning. Post-exploitation automation has also been studied using the same artificial intelligence techniques (Maeda & Mimura, 2020). These methods used in combination with the state-of-the-art tools would enable a cheaper price point to maintain a cyber range.

6 Conclusion & future work

In this paper, we summarized a list of takeaways and considerations for future researchers to improve cybercompetitions. We built a taxonomy of different cybercompetition types and related them to recognized cybersecurity disciplines, as explained by the InfoSEC color wheel. In doing so, we provided a table that broke down the expected learning outcomes from each type of competition. By evaluating the state-of-the-art technologies which cybercompetitions rely on, we hypothesized challenges technologies must resolve and provided analysis of changes which must be accomplished to host successful cybercompetitions that will attract students into the domain of cybersecurity.

During the course of our research in this systematization of knowledge, our lab hosted our own cybersecurity boot-camp to which we applied the knowledge gained from our noted takeaways. Subsequent papers will follow outlining the lessons learned from our boot-camp, notable changes we made in hosting our competition, feedback from our participants, and we will strategize future changes to continue to empower the next generation with skill development in cybersecurity.

References

Abomhara, M., et al. (2015). Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. *Journal of Cyber Security and Mobility*, *4*(1), 65–88. https://doi.org/10.13052/jcsm2245-1439.414.

Antonaci, A., Klemke, R., Stracke, C. M., Specht, M., Spatafora, M., & Stefanova, K. (2017). Gamification to empower information security education. In *International gamiFIN conference 2017* (pp. 32–38).

April, W. (2017). Orange is the new purple. *BlackHat USA*, *2017*, 1–9. https://www.blackhat.com/docs/us-17/wednesday/us-17-Wright-Orange-Is-The-New-Purple-wp.pdf?ref=hackernoon.com. Accessed 11 Jan 2023.

apsdehal (2020). awesome-ctf. https://github.com/apsdehal/awesome-ctfl. Accessed 11 Jan 2023.

Bashir, M., Lambert, A., Guo, B., Memon, N., & Halevi, T. (2015). Cybersecurity competitions: The human angle. *IEEE Security & Privacy*, *13*(5), 74–79. https://doi.org/10.1109/MSP.2015.100.

Benson, J. O., Prevost, J. J., & Rad, P. (2016). Survey of automated software deployment for computational and engineering research. In *2016 Annual IEEE systems conference (syscon)* (pp. 1–6). IEEE, https://doi.org/10.1109/SYSCON.2016.7490666.

Blohm, I., & Leimeister, J. M. (2013). Gamification. *Business & Information Systems Engineering*, *5*(4), 275–278.

Boopathi, K., Sreejith, S., & Bithin, A. (2015). Learning cyber security through gamification. *Indian Journal of Science and Technology*, *8*(7), 642–649. https://doi.org/10.17485/ijst/2015/v8i7/67760.

Brooks, T. T., Caicedo, C., & Park, J. S (2012). Security vulnerability analysis in virtualized computing environments. *International Journal of Intelligent Computing Research*, *3*(1/2), 277–291.

Brotherston, L., & Berlin, A. (2017). Defensive security handbook: best practices for securing infrastructure. " O'Reilly Media Inc.".

Bumanglag, K., Law, D., Welle, A., & Barrett, P. (2019). Constructing large scale cyber wargames. In *International conference on cyber warfare and security* (pp. 653–x). Academic conferences international limited.

Burket, J., Chapman, P., Becker, T., Ganas, C., & Brumley, D. (2015). Automatic problem generation for capture-the-flag competitions. In *2015 {USENIX} Summit on gaming, games, and gamification in security education (3GSE 15)*.

CTFTime (2020). About ctftim. https://ctftime.org/about/. Accessed 11 Jan 2023.

CTFWiki Team (2019). Ctf wiki - history of competitions. https://ctf-wiki.github.io/ctf-wiki/introduction/history/. Accessed 11 Jan 2023.

Carnegie Mellon University (2018). SEI. Vtunnel https://github.com/cmu-sei/vtunnel. Accessed 11 Jan 2023.
Carnegie Mellon University (2019). SEI. Welle-D https://github.com/cmu-sei/welled. Accessed 11 Jan 2023.
Carnegie Mellon University (2020a). picoctf. https://github.com/picoCTF/picoCTF. Accessed 11 Jan 2023.
Carnegie Mellon University (2020b). SEI. Topgen https://github.com/cmu-sei/topgen. Accessed 11 Jan 2023.
Carnegie Mellon University (2020c). SEI. Greybox https://github.com/cmu-sei/greybox. Accessed 11 Jan 2023.
Carnegie Mellon University (2020d). SEI. Ghosts Https://github.com/cmu-sei/GHOSTS. Accessed 11 Jan 2023.
Carnegie Mellon University (2020e). SEI. Topomojo https://github.com/cmu-sei/TopoMojo. Accessed 11 Jan 2023.
Center for Cyber Safety and Education (2019). Strategies for building and growing strong cybersecurity team. https://www.isc2.org/-/media/ISC2/Research/2019-Cybersecurity-Workforce-Study/ISC2-Cybersecurity-Workforce-Study-2019.ashx. Accessed 07 May 2020.
Chaisiri, S., Lee, B.-S., & Niyato, D. (2011). Optimization of resource provisioning cost in cloud computing. *IEEE Transactions on Services Computing*, *5*(2), 164–177. https://doi.org/10.1109/TSC.2011.7.
Chapman, P., & Brumley, D. (2013). picoctf: Teaching 10,000 high school students to hack. CMU.
Chapman, P., Burket, J., & Brumley, D. (2014). Picoctf: A game-based computer security competition for high school students. In *2014* {*USENIX*} *Summit on gaming, games, and gamification in security education (3GSE 14)*.
Cheung, R. S., Cohen, J. P., Lo, H. Z., Elia, F., & Carrillo-Marquez, V. (2012). Effectiveness of cybersecurity competitions. In *Proceedings of the International Conference on Security and Management (SAM). The Steering Committee of The World Congress in Computer Science, Computer* ... (p. 1).
Chu, B.-T., Ahn, G.-J., Blanchard, S., Deese, J., Kelly, R., Yu, H., & Young, A. (2007). Collegiate cyber game design criteria and participation. In *6Th IEEE/ACIS international conference on computer and information science (ICIS 2007)* (pp. 1036–1041). IEEE, https://doi.org/10.1109/ICIS.2007.80.
Chung, K. (2020a). Ctfd. https://github.com/CTFd/CTFd. Accessed 11 Jan 2023.
Chung, K. (2020b). Ctfd plugins https://github.com/CTFd/plugins. Accessed 11 Jan 2023.
Chung, K., & Cohen, J. (2014). Learning obstacles in the capture the flag model. In *2014* {*USENIX*} *Summit on gaming, games, and gamification in security education (3GSE 14)*.
Cisco Systems (2018). Annual internet report. https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/air-highlights.html#. Accessed 11 Jan 2023.
Conklin, A. (2006). Cyber defense competitions and information security education: An active learning solution for a capstone course. In *Proceedings of the 39th annual hawaii international conference on system sciences (HICSS'06)*, (Vol. 9 pp. 220b–220b). IEEE, https://doi.org/10.1109/HICSS.2006.110.
Conti, G., Babbitt, T., & Nelson, J. (2011). Hacking competitions and their untapped potential for security education. *IEEE Security & Privacy*, *9*(3), 56–59. https://doi.org/10.1109/MSP.2011.51.
Cowan, C., Arnold, S., Beattie, S., Wright, C., & Viega, J. (2003). Defcon capture the flag Defending vulnerable code from intense attack. In *Proceedings DARPA information survivability conference and exposition*, (Vol. 1 pp. 120–129). IEEE, https://doi.org/10.1109/DISCEX.2003.1194878.
Cremen, L. (2020). Introducing the infosec colour wheel – blending developers with red and blue security teams. https://tinyurl.com/wuuvnfl. Accessed 07 May 2020.
Cyberbit (2018). Cyber range for higher education - build or buy? Technical report, Cyberbit. https://storage.googleapis.com/stateless-www-cyberbit-com-liv/2018/10/Cyber-Range-Build-vs.-Buy-White-Paper.pdf.
DEFCON (2018). A history of capture the flag at def con. https://www.defcon.org/html/links/dc-ctf-history.html. Accessed 11 Jan 2023.
Dabrowski, A., Kammerstetter, M., Thamm, E., Weippl, E., & Kastner, W. (2015). Leveraging competitive gamification for sustainable fun and profit in security education. In *2015* {*USENIX*} *Summit on gaming, games, and gamification in security education (3GSE 15)*.
Department of Homeland Security (2020). Fact sheet: Executive order on cybersecurity / presidential policy directive on critical infrastructure security and resilience. https://www.dhs.gov/news/2013/02/13/fact-sheet-executive-order-cybersecurity-presidential-policy-directive-critical. Accessed 07 May 2020.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining" gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments.* (pp. 9–15), https://doi.org/10.1145/2181037.2181040.

Dillon, T., Wu, C., & Chang, E. (2010). Cloud computing: issues and challenges. In *2010 24Th IEEE international conference on advanced information networking and applications* (pp. 27–33). Ieee, https://doi.org/10.1109/AINA.2010.187.

Dodge, R. C., & Ragsdale, D. J. (2004). Organized cyber defense competitions. In *IEEE International conference on advanced learning technologies, 2004. Proceedings* (pp. 768–770). IEEE, https://doi.org/10.1109/ICALT.2004.1357651.

Eagle, C. (2013). Computer security competitions. Expanding educational outcomes. *IEEE Security & Privacy*, *11*(4), 69–71. https://doi.org/10.1109/MSP.2013.83.

Eagle, C., & Clark, J. L. (2004). Capture-the-flag Learning computer security under fire. Technical report, Naval postgraduate school MONTEREY CA. https://apps.dtic.mil/sti/citations/ADA435319. Accessed 11 Jan 2023.

EasyCTF (2019). Librectf https://github.com/easyctf/librectf. Accessed 11 Jan 2023.

Echothrust Solutions (2020). echoctf. https://github.com/echoCTF/echoCTF.RED. Accessed 11 Jan 2023.

Evans, K., & Reeder, F. (2010). A human capital crisis in cybersecurity Technical proficiency matters. CSIS.

Facebook (2018). Fbctf. https://github.com/facebookarchive/fbctf. Accessed 11 Jan 2023.

Ferguson, B., Tall, A., & Olsen, D. (2014). National cyber range overview. In *2014 IEEE Military communications conference* (pp. 123–128), https://doi.org/10.1109/MILCOM.2014.27.

Fink, G., Best, D., Manz, D., Popovsky, V., & Endicott-Popovsky, B. (2013). Gamification for measuring cyber security situational awareness. In *International conference on augmented cognition.* (pp. 656–665). Springer, https://doi.org/10.1007/978-3-642-39454-6_70.

Fouché, S., & Mangle, A. H. (2015). Code hunt as platform for gamification of cybersecurity training. In *Proceedings of the 1st international workshop on code hunt workshop on educational software engineering.* (pp. 9–11), https://doi.org/10.1145/2792404.2792406.

Gavas, E., Memon, N., & Britton, D. (2012). Winning cybersecurity one challenge at a time. *IEEE Security & Privacy*, *10*(4), 75–79. https://doi.org/10.1109/MSP.2012.112.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?–a literature review of empirical studies on gamification. In *2014 47Th hawaii international conference on system sciences.* (pp. 3025–3034). Ieee, https://doi.org/10.1109/HICSS.2014.377.

Hassenfeldt, C., Jacques, J., & Baggili, I. (2020). Exploring the learning efficacy of digital forensics concepts and bagging & tagging of digital devices in immersive virtual reality. *Forensic Science International: Digital Investigation*, *33*, 301011. https://doi.org/10.1016/j.fsidi.2020.301011.

Hayes, B. (2008). Cloud computing. https://doi.org/10.1145/1364782.1364786.

Hochstein, L., & Moser, R. (2017). Ansible: Up and running: automating configuration management and deployment the easy way. " O'Reilly Media Inc.".

Kafai, Y. B., & Burke, Q. (2014). *Connected code: Why children need to learn programming*. Cambridge: Mit Press.

Kevin, C. (2018). Ctfd-docker https://github.com/CTFd/CTFd-docker. Accessed 11 Jan 2023.

Kianpour, M., Kowalski, S., Zoto, E., Frantz, C., & Øverby, H. (2019). Designing serious games for cyber ranges A socio-technical approach. In *2019 IEEE European symposium on security and privacy workshops (euros PW)* (pp. 85–93), https://doi.org/10.1109/EuroSPW.2019.00016.

Kocher, P., Horn, J., Fogh, A., Genkin, D., Gruss, D., Haas, W., Hamburg, M., Lipp, M., Mangard, S., Prescher, T., Schwarz, M., & Yarom, Y. (2019). Spectre attacks: Exploiting speculative execution. In *40Th IEEE symposium on security and privacy (s&p'19).* https://doi.org/doi.org/10.1145/3399742.

Kolias, C., Kambourakis, G., Stavrou, A., & Voas, J. (2017). Ddos in the iot: Mirai and other botnets. *Computer*, *50*(7), 80–84. https://doi.org/10.1109/MC.2017.201.

Ladabouche, T., & LaFountain, S. (2016). Gencyber Inspiring the next generation of cyber stars. *IEEE Security Privacy*, *14*(5), 84–86. https://www.doi.org/10.1109/MSP.2016.107.

Langevin, R. J. R., McCaul, R. M. T., Charney, S., Lt, G. H. R., & Lewis, J.A. (2011). Cybersecurity two years later: Washington DC: Center for Strategic and International Studies. http://static.cs.brown.edu/courses/csci1800/sources/2011_CSIS_Lewis_CybersecurityTwoYearsLater.pdf. Accessed 11 Jan 2023.

Langevin, J. R., McCaul, M. T., Charney, S., & Raduege, H. (2008). Securing cyberspace for the 44th presidency. Technical report: Center for strategic and international studies WASHINGTON DC. https://apps.dtic.mil/sti/citations/ADA490800. Accessed 11 Jan 2023.

Lei, K., Ma, Y., & Tan, Z. (2014). Performance comparison and evaluation of web development technologies in php, python, and node. js. In *2014 IEEE 17Th international conference on computational science and engineering*. https://doi.org/10.1109/CSE.2014.142 (pp. 661–668). IEEE.

Leune, K., & Petrilli, S. J. Jr. (2017). Using capture-the-flag to enhance the effectiveness of cybersecurity education. In *Proceedings of the 18th annual conference on information technology education* (pp. 47–52), https://doi.org/10.1145/3125659.3125686.

Levinson, A. (2019a). Deep exploit https://github.com/13o-bbr-bbq/machine_learning_security/tree/master/De. Accessed 11 Jan 2023.

Levinson, A. (2020b). laforge. https://github.com/gen0cide/laforge. Accessed 11 Jan 2023.

Li, Y., Nguyen, D., & Xie, M. (2017). Ezsetup: A novel tool for cybersecurity practices utilizing cloud resources. In *Proceedings of the 18th annual conference on information technology education* (pp. 53–58), https://doi.org/10.1145/3125659.3125699.

Lipp, M., Schwarz, M., Gruss, D., Prescher, T., Haas, W., Fogh, A., Horn, J., Mangard, S., Kocher, P., Genkin, D., Yarom, Y., & Hamburg, M. (2018). Meltdown: Reading kernel memory from user space. In *27Th USENIX security symposium (USENIX security 18)*.

MITRE (2018). cascade-server. https://github.com/mitre/cascade-server. Accessed 11 Jan 2023.

MITRE (2020a). Mite ctf scoreboard. https://github.com/mitre-cyber-academy/ctf-scoreboard. Accessed 11 Jan 2023.

MITRE (2020b). caldera. https://github.com/mitre/caldera. Accessed 11 Jan 2023.

MITRE (2020c). car. https://github.com/mitre/car. Accessed 11 Jan 2023.

Maeda, R., & Mimura, M. (2020). Automating post-exploitation with deep reinforcement learning. Computers & Security p. 102108 https://doi.org/10.1016/j.cose.2020.102108.

Martens, B., & Teuteberg, F. (2012). Decision-making in cloud computing environments A cost and risk based approach. *Information Systems Frontiers*, *14*(4), 871–893. https://doi.org/10.1007/s10796-011-9317-x.

Matias, P., Barbosa, P., Cardoso, T. N. C., Campos, D. M., & Aranha, D.F. (2018). Nizkctf: A noninteractive zero-knowledge capture-the-flag platform. *IEEE Security & Privacy*, *16*(6), 42–51. https://doi.org/10.1109/MSEC.2018.2875324.

Mayes, J. (2020). Six free tools for creating a cyber simulator. https://insights.sei.cmu.edu/sei_blog/2019/04/six-free-tools-for-creating-a-cyber-simulator.html. Accessed 07 May 2020.

Miessler, D. (2020). The difference between red, blue, and purple teams. https://tinyurl.com/shueqjg. Accessed 07 May 2020.

Morris, L. (2020). Nationalcptc - report examples https://github.com/nationalcptc/report_examples. Accessed 11 Jan 2023.

Moscovich, N., Bitton, R., Mallah, Y., Inokuchi, M., Yagyu, T., Kalech, M., Elovici, Y., & Shabtai, A. (2020). Autosploit: A fully automated framework for evaluating the exploitability of security vulnerabilities. arXiv:2007.00059.

Moss, J. (2008). Off at a tangent—a discussion with jeff moss. *Computer Fraud & Security*, *2008*(9), 7–10. https://doi.org/10.1016/S1361-3723(08)70136-2.

Munaiah, N., Pelletier, J., Su, S.-H., Yang, J. S., & Meneely, A. (2019). A cybersecurity dataset derived from the national collegiate penetration testing competition. In *HICSS Symposium on cybersecurity big data analytics*.

National Institute of Standards US Department of Commerce and Technology (2020). Executive order on america's cybersecurity workforce. https://www.nist.gov/news-events/news/2019/05/executive-order-americas-cybersecurity-workforce. Accessed 07 May 2020.

National Institute of Standards and Technology (2020). Nist: Cyber ranges. https://www.nist.gov/system/files/documents/2018/02/13/cyber_ranges.pdf. Accessed 28 Nov 2020.

Ncac - faq (2019). https://cyberanalystchallenge.org/faq/.

Newby, C. R. (2018). Designing cybersecurity competitions in the cloud: A framework and feasibility study. https://www.proquest.com/dissertationstheses/designing-cybersecurity-competitionscloud/docview/2439000969/se-2.

Nguyen-Tuong, A., Melski, D., Davidson, J. W., Co, M., Hawkins, W., Hiser, J. D., Morris, D., Nguyen, D., & Rizzi, E. (2018). Xandra: An autonomous cyber battle system for the cyber grand challenge. *IEEE Security Privacy*, *16*(2), 42–51. https://doi.org/10.1109/MSP.2018.1870876.

Office of the Press Secretary (2020a). President donald j. trump is strengthening america's cybersecurity workforce to secure our nation and promote prosperity. https://www.whitehouse.gov/briefings-statements/president-donald-j-trum p-strengthening-americas-cybersecurity-workforce-secure-nation-promote-prosper ity/. Accessed 07 May 2020.

Office of the Press Secretary (2020b). Fact sheet: Cybersecurity national action plan. https://obamawhitehouse.archives.gov/the-press-office/2016/02/09/fact-sheet-cybersecurity-national-action-plan. Accessed 07 May 2020.

Owens, K., Fulton, A., Jones, L., & Carlisle, M. (2020). Pico-boo!: How to avoid scaring students away in a ctf competition.

Patriciu, V.-V., & Furtuna, A. C. (2009). Guide for designing cyber security exercises. In *Proceedings of the 8th WSEAS international conference on e-activities and information security and privacy* (pp. 172–177). World Scientific and Engineering Academy and Society (WSEAS).

Paulsen, C., McDuffie, E., Newhouse, W., & Toth, P. (2012). Nice: Creating a cybersecurity workforce and aware public. *IEEE Security & Privacy*, *10*(3), 76–79. https://doi.org/10.1109/MSP.2012.73.

Peter, M., Grance, T., & et al (2011). The nist definition of cloud computing. http://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf. Accessed 11 Jan 2023.

Petersen, R., & Santos, D. (2020). (manager of communications and operations) karen a. wetzel (manager of the nice framework) national initiative for cybersecurity education (nice) applied cybersecurity division information technology laboratory. *NIST Special Publication*, *800*, 181. https://doi.org/10.6028/NIST.SP.800-181r1.

Popović, K., & Željko, H. (2010). Cloud computing security issues and challenges. In *The 33rd international convention mipro* (pp. 344–349). IEEE.

Resnick, M. (2013). Learn to code, code to learn EdSurge, May 54 https://el.media.mit.edu/logo-foundation/services/pdf/program2013.pdf. Accessed 11 Jan 2023.

Ruan, K., Baggili, I., Carthy, J., & Kechadi, T. (2011). Survey on cloud forensics and critical criteria for cloud forensic capability: A preliminary analysis. https://doi.org/10.1016/j.diin.2013.02.004.

Sabbath, J. (2020). Virginia cyber range to grow under new agreement. https://www.virginiabusiness.com/article/virginia-cyber-range-to-grow-under-new-agreement/. Accessed 28 Nov 2020.

Sahoo, J., Mohapatra, S., & Lath, R. (2010). Virtualization: A survey on concepts, taxonomy and associated security issues. In *2010 Second international conference on computer and network technology* (pp. 222–226). IEEE, https://doi.org/10.1109/ICCNT.2010.49.

Sajid, M., & Raza, Z. (2013). Cloud computing: Issues & challenges. In *International conference on cloud, big data and trust*, (Vol. 20 pp. 13–15).

Saman, A. A. (2007). Observations on teamwork strategies in the acm international collegiate programming contest. XRDS: Crossroads. *The ACM Magazine for Students*, *14*(1), 1–9.

Scholefield, S., & Shepherd, L. A. (2019). Gamification techniques for raising cyber security awareness. In *International conference on human-computer interaction* (pp. 191–203). Springer, https://doi.org/10.1007/978-3-030-22351-9_13.

Schreuders, C. Z., & Butterfield, E. (2016). Gamification for teaching and learning computer security in higher education. In *2016 {USENIX} Workshop on advances in security education ({ASE} 16)*.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006.

Shein, E. (2014). Should everybody learn to code? https://doi.org/10.1145/2557447.

Singh, G. (2020). Facebook ctf is now open source!. https://www.facebook.com/notes/facebook-ctf/facebook-ctf-is-now-open-source/525464774322241. Accessed 11 Jan 2023.

Sommestad, T., & Hallberg, J. (2012). Cyber security exercises and competitions as a platform for cyber security experiments. In *Nordic conference on secure IT systems.* https://doi.org/10.1007/978-3-642-34210-3_4 (pp. 47–60). Springer.

Song, J., & Alves-Foss, J. (2015). The darpa cyber grand challenge: A competitor's perspective. *IEEE Security Privacy*, *13*(6), 72–76. https://doi.org/10.1109/MSP.2015.132.

Song, J., & Alves-Foss, J. (2016). The darpa cyber grand challenge: a competitor's perspective, part 2. *IEEE Security Privacy*, *14*(1), 76–81. https://doi.org/10.1109/MSP.2016.14.

Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C.B. (2018). att&ck Mitre Design and philosophy Technical report.

Taylor, C., Arias, P., Klopchic, J., Matarazzo, C., & Evi, D. (2017). {CTF}: State-Of-the-art and building the next generation. In *2017 {USENIX} Workshop on advances in security education ({ASE} 17)*.

Temple University (2020). Collegiate social engineering competition and training event. https://sites.temple.edu/socialengineering/. Accessed 11 Jan 2023.
Texas, A.&M. (2020). Portable challenges plugin https://github.com/tamuctf/ctfd-portable-challenges-plugin. Accessed 11 Jan 2023.
Thornton, D., & Francia, G. (2014). Gamification of information systems and security training: Issues and case studies. *Information Security Education Journal*, *1*(1), 16–24. https://www.dline.info/isej/fulltext/v1n1/3.pdf. Accessed 11 Jan 2023.
Tobey, D. H., Pusey, P., & Burley, D.L. (2014). Engaging learners in cybersecurity careers: lessons from the launch of the national cyber league. *ACM Inroads*, *5*(1), 53–56. https://doi.org/10.1145/2568195.2568213.
Tyler, J. (2018). Hash crack king plugin https://github.com/akm0d/CTFd_hash_crack_king. Accessed 11 Jan 2023.
US Department of Defense (2019). Contracts for nov. 20. https://www.defense.gov/Newsroom/Contracts/Contract/Article/2022193/#LOCKHEED112019. Accessed 28 Nov 2020.
University of Connecticut (2020). Csi cyberseed. https://csi.uconn.edu/cybersecurity-week-2017/#. Accessed 14 May 2020.
UnrealAkama (avidhacker) (2017). Nightshade https://github.com/UnrealAkama/NightShade. Accessed 11 Jan 2023.
Van der Vegt, W. (2006). The codecup, an annual game programming competition Perspectives on computer science competitions for (high school) students https://ioi.te.lv/workshop/dagstuhl_2006/RevisedPapers/3_vanderVegt_rev.pdf. Accessed 11 Jan 2023.
Weisse, O., Van Bulck, J., Minkin, M., Genkin, D., Kasikci, B., Piessens, F., Silberstein, M., Strackx, R., Wenisch, T. F., & Yarom, Y. (2018). Foreshadow-ng: Breaking the virtual memory abstraction with transient out-of-order execution.
White, G. B., Williams, D., & Harrison, K. (2010). The cyberpatriot national high school cyber defense competition. *IEEE Security & Privacy*, *8*(5), 59–61. https://doi.org/10.1109/MSP.2010.166.
Williams, D. (2020). Ccdc - 2020 rules. https://tinyurl.com/ydxf4mcu. Accessed 07 May 2020.
Woszczynski, A. B., & Green, A. (2017). Learning outcomes for cyber defense competitions. *Journal of Information Systems Education*, *28*(1), 21.
Xie, T., Bishop, J., Tillmann, N., & Halleux, J. D. (2015). Gamifying software security education and training via secure coding duels in code hunt. In *Proceedings of the 2015 symposium and bootcamp on the science of security*. https://doi.org/10.1145/2746194.2746220 (pp. 1–2).
XuCcc (2018). Ctfdonlinechallenge https://github.com/xuccc/CTFdonlinechallenge. Accessed 11 Jan 2023.
Yevgeniy, B. (2016). Why we use terraform and not chef, puppet, ansible, saltstack or cloudformation https://lsi.vc.ehu.eus/pablogn/docencia/AS/Aact7%20Admin.%20centralizada/Terraform%20Chef%20Puppet20Ansible%20Salt.pdf. Accessed 11 Nov 2023.
Yonemura, K., Yajima, K., Komura, R., Sato, J., & Takeichi, Y. (2017). Practical security education on operational technology using gamification method. In *2017 7Th IEEE international conference on control system, computing and engineering (ICCSCE)*. https://doi.org/10.1109/ICCSCE.2017.8284420 (pp. 284–288). IEEE.
Yu, L.-H. (2016). Rc3-ctf-2016-scoreboard https://github.com/seadog007/RC3-CTF-2016-scoreboard. Accessed 11 Jan 2023.
Zhang, K., Dong, S., Zhu, G., Corporon, D., McMullan, T., & Barrera, S. (2013). Picoctf 2013-toaster wars: When interactive storytelling game meets the largest computer security competition. In *2013 IEEE International games innovation conference (IGIC)* (pp. 293–299). IEEE.
Zichermann, G., & Cunningham, C. (2011). Gamification by design: Implementing game mechanics in web and mobile apps. " O'Reilly Media Inc.".