Lessons Learned From Complex Hands-on Defence Exercises in a Cyber Range

Jan Vykopal*, Martin Vizvary*, Radek Oslejsek†, Pavel Celeda* and Daniel Tovarnak*
*Masaryk University, Institute of Computer Science, Brno, Czech Republic
†Masaryk University, Faculty of Informatics, Brno, Czech Republic
Emails: {vykopal|vizvary|celeda|tovarnak}@ics.muni.cz, oslejsek@fi.muni.cz

***Abstract*—We need more skilled cybersecurity professionals because the number of cyber threats and ingenuity of attackers is ever growing. Knowledge and skills required for cyber defence can be developed and exercised by lectures and lab sessions, or by active learning, which is seen as a promising and attractive alternative. In this paper, we present experience gained from the preparation and execution of cyber defence exercises involving various participants in a cyber range. The exercises follow a Red vs. Blue team format, in which the Red team conducts malicious activities against emulated networks and systems that have to be defended by Blue teams of learners. Although this exercise format is popular and used worldwide by numerous organizers in practice, it has been sparsely researched. We contribute to the topic by describing the general exercise life cycle, covering the exercise's development, dry run, execution, evaluation, and repetition. Each phase brings several challenges that exercise organizers have to deal with. We present lessons learned that can help organizers to prepare, run and repeat successful events systematically, with lower effort and costs, and avoid a trial-and-error approach that is often used.**

I. Introduction

Information and communication systems are exposed to an increasing number of attacks. Apart from simple attacks conducted by hacktivists and inexperienced individuals that can be tracked down [1], there are professional teams backed by organized crime groups or even governments [2] that carefully hide their activities. A shortage in cyber security skills and cyber security professionals is a critical vulnerability for companies and nations [3], [4].

Cyber security can be taught not only using conventional methods, including classroom lectures, seminars or home assignments, but also by hands-on experience. In recent years, there has been a significant growth of hands-on competitions, challenges, and exercises [5], [6]. It is believed that they enable participants to effectively gain or practise diverse cyber security skills in an attractive way.

The most popular events are Capture The Flag (CTF) games [5] and Cyber Defence eXercises (CDX) [6]. While CTF games focus on attacking, defending or both, CDXs train solely the defence. CTFs which put participants in the role of the attacker support the development of adversarial thinking that is necessary for anticipating future offensive actions [7]. CDXs enable participants to experience cyber attacks first-hand.

Although both types of events are prepared and carried out by numerous sponsors for a large number of participants, there are only a few public research papers dealing with the design of an exercise in a cyber range. Granåsen and Andersson conducted a case study on measuring team effectiveness in Baltic Cyber Shield 2010, a multi-national civil-military CDX [8]. They described the instrumentation and collection of data from the exercise's infrastructure and participants in order to provide situational awareness for organizers during the exercise. The Spanish National Cybersecurity Institute proposed a taxonomy of cyber exercises [9] which recognizes operations-based exercises focused on incident response by participants in technical and management roles. ISO/TC 223 effort resulted into ISO 22398, which describe general guidelines for exercises including basic terms and definitions [10]. Unfortunately, technical implementation details of an exercise in a cyber range is out of scope of this standard.

In our work, we address the gap in the literature by describing the life cycle of a complex cyber defence exercise and challenges related to the exercise's design, development, execution and repeatability. This knowledge is based on our experience gained by developing and delivering six runs of a cyber defence exercise scenario with about 120 national and international learners between 2015 and 2017. The exercises have been carried out in a cyber range we are developing and continuously enhancing in order to suit an exercise's requirements.

This paper is organized into five sections. Section 2 provides an overview of existing platforms that can be used as a vehicle for cyber exercises. Section 3 describes a cyber defence exercise carried out in a cyber range. Section 4 reports on lessons learned through six runs of this exercise. Finally, Section 5 concludes the paper and outlines future work.

II. Hands-on learning environments

In this section, we give a brief overview of learning environments that can be used in active learning of cyber security. We have done a systematic literature review from 2013 to 2017 to cover recent advances and innovations.

A. Generic testbeds

Generic testbeds provide a basic functionality for the emulation of computer networks. *Emulab/Netbed* [11] is

978-1-5090-5920-1/17/$31.00 ©2017 IEEE

a cluster testbed providing services for the deployment of virtual appliances, configuration of flexible network topologies and emulation of various network characteristics. Emulab allocates computing resources for a specified network and instantiates it at a dedicated hardware infrastructure. *CyberVAN* [12] experimentation testbed provides a virtualized environment where arbitrary applications running on Xen-based virtual machines can be interconnected by arbitrary network topologies. It employs network simulators such as OPNET, QualNet, ns-2, or ns-3, so the network traffic of emulated hosts travels through the simulated network. This hybrid emulation enables the simulation of large strategic networks approximating a large ISP network.

B. Lightweight platforms

Several lightweight platforms have been developed for cyber security training. While some of them evolved from the generic testbeds, others were designed from scratch with different needs in mind. *Avatao* [13], [14] is a web-based online e-learning platform offering IT security challenges (hands-on exercises), which can be organized to a path which leads to fulfilling an ultimate learning objective. *CTF365* [15] (Capture The Flag 365) is a training platform that leverages gamification to improve retention rate and speed up the learning and training curve. In the *Hacking-Lab* [16] online platform, teams of participants have to perform several tasks simultaneously; keep applications up and running, find and patch vulnerabilities, solve challenges and attack their competitors' applications.The *iCTF* framework [17] has been developed at The University of California for hosting their iCTF, the largest capture the flag competition in the world. *InCTF* [18] is a modification of the iCTF framework. Using Docker containers instead of virtual machines enhances the overall game experience and simplifies the organization of attack-defence competitions for a larger number of participants.

C. Cyber ranges

Cyber ranges represent complex virtual environments that are used not only for cyberwarfare training, but also for cybertechnology development, forensic analysis and other cyber-related issues. There is an extensive survey of state-of-the-art cyber ranges and testbeds [19]. One very popular cyber range is *DETER/DeterLab* [20], [21], which is based on Emulab and was started with the goal of advancing cyber security research and education in 2004. Nowadays, there exist many other cyber ranges, e. g., *National Cyber Range (NCR)* [22], *Michigan Cyber Range (MCR)* [23], *SimSpace Cyber Range* [24], *EDURange* [25], or *KYPO Cyber Range* [26].

III. Cyber defence exercise

We have designed a one day Red vs. Blue cyber defence exercise for 50 participants. It was inspired by the Locked Shield exercise [27] organized by NATO Cooperative Cyber Defence Centre of Excellence in Tallinn. We named our exercise Cyber Czech and it has been executed six times so far (2015–2017). Cyber Czech is a hands-on exercise improving the technical and soft skills of security professionals grouped in six Blue teams. It requires substantial preparation effort from the organizers and a dedicated cyber range infrastructure. The exercise involves:

- cloud-based exercise infrastructure (sandboxes),
- training objectives, story, and an exercise scenario,
- participants grouped in teams (Red, Blue, White and Green),
- a physical cyber range facility hosting all participants.

This section explains the cyber defence exercise's components, terms used and definitions, we will use throughout the rest of the paper.

A. Cyber range infrastructure

The technical part of the cyber exercise relies on a cyber range itself and supportive infrastructure for communication within the exercise and the evaluation of participants' actions. The cyber range emulates a complex network setup in a contained environment. Therefore, participants can realistically interact with an assigned host or network infrastructure, and their actions cannot interfere with the operational environment. The following text describes a high-level view of the architecture of the KYPO [26] cyber range, which we use in the Cyber Czech exercise.

Sandboxes represent a low-level layer of the cyber range. They encapsulate isolated computer networks where users can safely perform their cyber security tasks. Sandboxes are based on virtual appliances placed in a cloud, which makes their allocation, replication, and maintenance easy. Despite the virtualization, neither users nor running applications can recognize that they do not run on a real network.

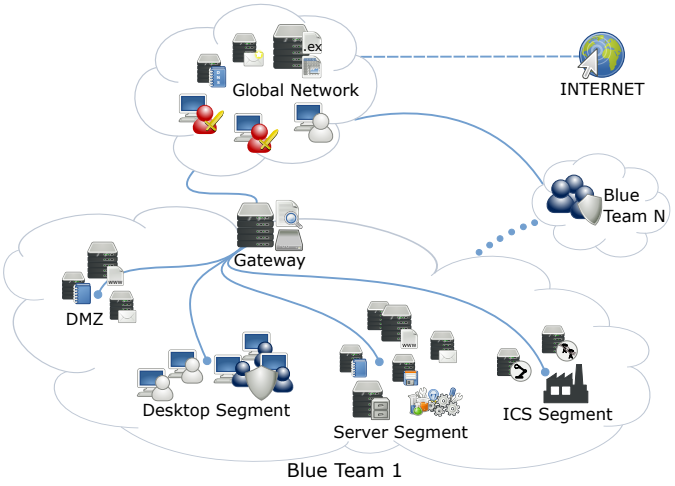

Figure 1. The scheme of the cyber exercise network.

The scheme of the cyber exercise network is depicted in Figure 1. This network serves as a virtual battlefield with approximately 110 interconnected hosts and other

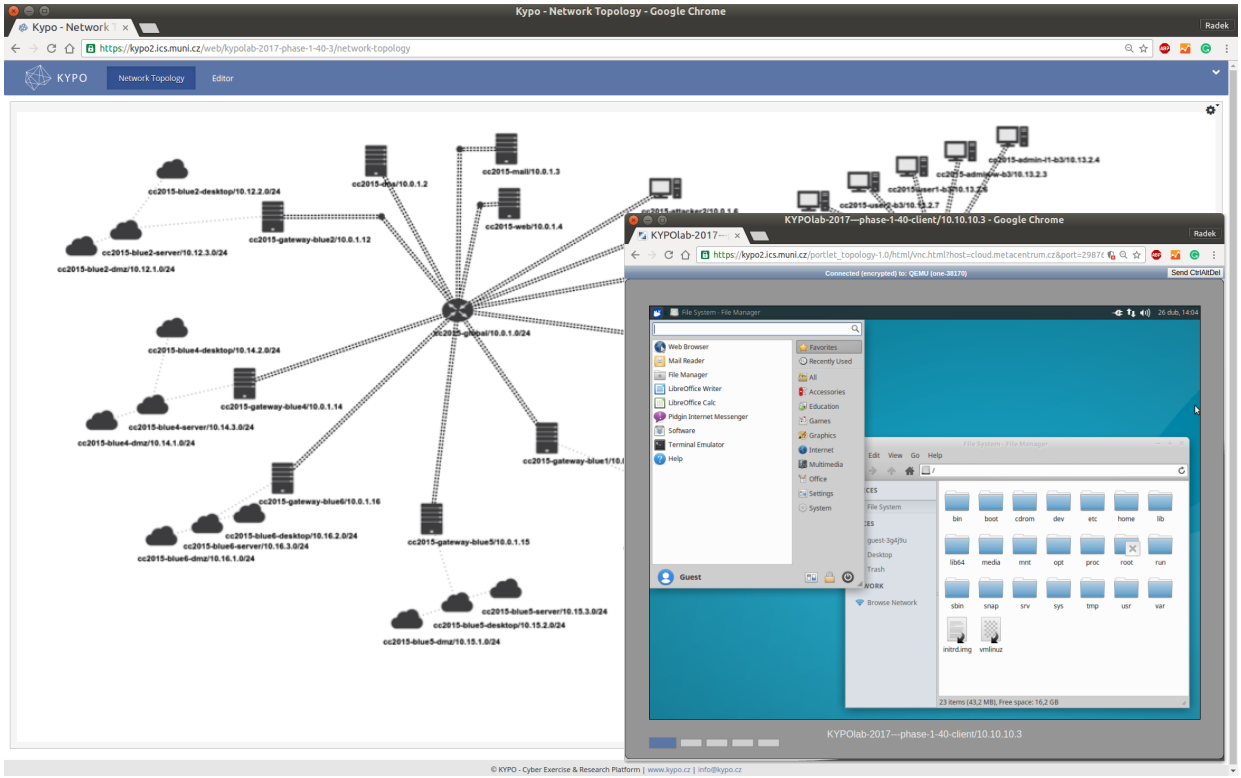

Figure 2. The topology of the exercise network, as seen by participants in the front-end application, and open remote desktop connection to selected host in the separate window.

network facilities. It is divided into two subnetworks: *i)* a global network hosting attackers and common network infrastructure, such as DNS and e-mail servers; this network simulates the global Internet, and *ii)* the networks of Blue teams representing the defended network with critical and vulnerable services. Networks of Blue teams are further divided into a demilitarized zone (DMZ), desktops, servers, and industrial control systems (ICS).

Cyber range built-in *monitoring services* cover network traffic statistics, flow data, and full-packet capture. In addition to these off-the-shelf data monitoring services, learners may install their own monitoring applications as a part of their activities inside their sandbox.

Next, we use a generic *logging infrastructure* integrated into the monitoring services. Each host is configured to forward log messages to the central logging server. A processing chain of additional tools is deployed in order to provide real-time access to the normalized log data from the exercise infrastructure. The state of the host's network services is periodically checked and events related to service state changes are logged into the central logging server.

The logging infrastructure is used by a *scoring system* that has been developed to provide feedback to participants during exercise. Penalty points are either computed automatically from events processed by the logging infrastructure (e. g., penalty for inaccessible services) or entered manually. A total score can be shown to participants in real-time. Monitored and logged data is an invaluable input for exercise management, evaluation and further research.

The *front-end application* provides a web-based user interface to interact with the cyber range. The web interface supports the design and management of sandboxes, single sign-on, remote desktop access etc. We have designed complex interactive visualizations to provide real-time feedback to participants, to provide insights into adversary behaviour, and to build effective situational awareness. Figure 2 shows a screenshot of a sandbox from the Cyber Czech exercise, as was seen by participants in the front-end application.

B. Exercise objectives, story and scenario

The designed exercise is focused on defending critical information infrastructure against skilled and coordinated attackers. Similarly to other defence exercises, learners are put into the role of members of emergency security teams which are sent into organizations to recover compromised networks. They have to secure the IT infrastructure, investigate possible data exfiltration and collaborate with other emergency teams, the coordinator of the operation and media representatives.

Learners are provided with a background story to introduce them to the situation before they enter the compromised networks. This is very important since the exercise is

not set in a real environment and learners have no previous knowledge who is who in the fictitious scenario (e. g., users in their organization, popular news portal, superordinate security team). They are also provided with technical facts related to the exercise network: network topology including "their" network that will be defended, network architecture and current setup, and access credentials, etc. Before the actual exercise, learners access their emulated network for several hours to get familiar with the exercise. The exercise is driven by a scenario which includes the actions of attackers and assignments for defenders prepared by the organizers. The attackers exploit specific vulnerabilities left in the compromised network in a fixed order which follows a common life cycle in the critical information infrastructure (see Figure 3).

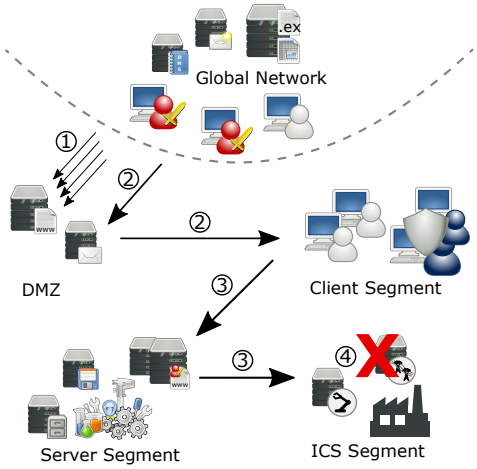

Figure 3. Common attack phases: ① reconnaissance the victim's network; ② exploitation of the unveiled vulnerabilities; ③ escalation of privileges on compromised computers and further exploitation; ④ completing attackers' mission, e. g., shutdown a control system.

The first attack phase involves reconnaissance (scanning of active systems or open network ports). Next, the attackers try to gain access to the machines providing public services (exploitation phase). This is followed by multiple escalations of privileges (accessing segments with internal machines), which enables the completing attackers' mission (shutdown of a critical application). The attackers use a mix of recent and ubiquitous attacks/vulnerabilities that are public and well-known. This is complemented by special tailored malware samples which emulate sophisticated attacks. The completion of each successful attack is recorded by the attackers. On top of that, learners should also answer media requests. The performance of each learners' team is scored based on successful attacks or their mitigation, the availability of specified critical services and the quality of reporting and communication.

C. Participant roles

Participants are divided into four groups according to their skills, role, and tasks in the exercise. These are now listed according to those commonly used in other cyber exercises:

- *Green team* – a group of operators responsible for the exercise infrastructure (the sandbox in this case). They configure all virtual computers and networks, monitoring and scoring infrastructure. The Green team also monitors the sandbox's health and fixes crashes and infrastructure issues if needed.
- *White team* – exercise managers, referees, organizers, and instructors. They provide the background story, exercise rules and framework for the Red team and Blue teams' competition. The White team assigns tasks (called injects) to the Blue teams and thus simulates media, the operation coordinator, and law enforcement agencies. They might also act as instructors and provide basic hints to Blue teams if needed.
- *Red team* – plays the role of attackers and consists of cyber security professionals. They do not attack targets in the infrastructure of a Blue team randomly, but carefully follow a predefined attack scenario to equally load the Blue teams. This means the Red team exploits vulnerabilities left in a Blue team's network. They should not use any other arbitrary means of attack against the Blue teams. They are also not allowed to attack the service infrastructure. Based on the success of attacks, the Red team assigns penalties to Blue teams. Penalties are assigned manually via a web interface since the amount of awarded points is based on non-trivial factors that need expert review.
- *Blue team* – learners responsible for securing compromised networks and dealing with the Red team's attacks. They have to follow the exercise's rules and local cyber law. The learners are grouped in several Blue teams.

Interactions between the four groups of participants are depicted in Figure 4.

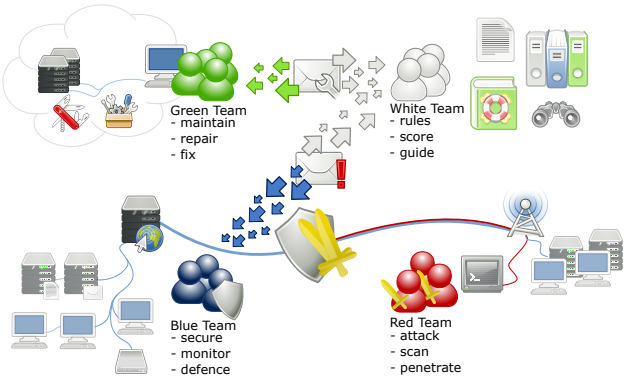

Figure 4. Exercise participants, their interactions and tasks.

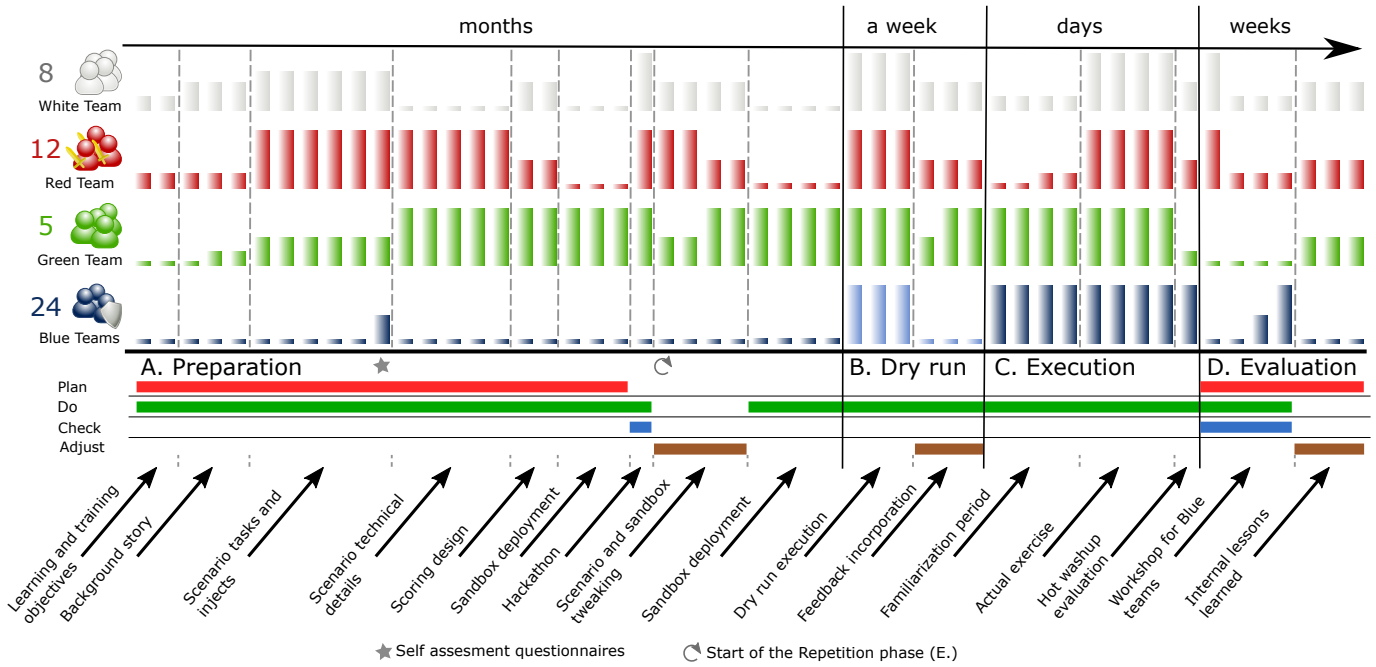

Figure 5. Cyber exercise life cycle in time. Coloured bars show relative effort spent by members of White, Red, Green and Blue teams in respective phases of the life cycle. The four numbers on the left express the size of particular team in the exercise. The mapping to the PDCA cycle is depicted by coloured lines below the life cycle phases.

IV. Lessons learned

Cyber exercises last several hours or days but their preparation typically takes several months involving experts from various fields – IT administrators, penetration testers, incident handlers, managers, legal experts etc. The exercise life cycle consists of several phases that can be mapped to a *Plan–Do–Check–Adjust* (PDCA) cycle. Carefully planning and considering the relationship of all phases may save a significant amount of invested effort and costs. Figure 5 shows the involvement of all teams and effort spent through the cyber exercise life cycle.

A. Preparation

The preparation phase consumes the majority of work effort and time. First, we have to set the learning and training objectives of the exercise; elaborate the background story and develop an exercise scenario consisting of tasks and injects for the Red team and White team – including end users, media and legal representatives. An outline of the exercise scenario is then used for preparing network infrastructure that will be defended by the Blue teams. A more detailed scenario is then used for setting up scoring components: their general weights (e. g., service availability vs. successful attacks vs. reporting) and score structure for every particular service, attack, or inject (e. g., if the Red team is successful in a given attack, the Blue team will be penalized by an exact number of points; if the Red team was successful only partially, the Blue team will be penalized only by a portion of the amount of full points).

In parallel, learners are invited and asked for self-assessment of their skills relevant to the exercise. Based on their input, the White team starts to create Blue teams with balanced skills and experience. The described steps so far correspond to the *Plan* and *Do* phases in the PDCA cycle.

Once the network infrastructure and hosts are configured according to the proposed scenario, they are deployed in a cyber range sandbox. Tasks and injects of the scenario are tested by members of Red team and White team in an intensive full day workshop (hackathon). This is without the presence of Blue teams. The hackathon represents the *Do* and the *Check* phases of the PDCA cycle. After that, there is the last chance to modify the scenario and configuration of exercise infrastructure (the *Adjust* phase).

In our experience, the most challenging tasks in the preparation phase are:

- *Setting learning objectives with respect to the expected readiness of prospective learners* – the organizers have limited information about learners' skills before the actual exercise. This is a completely different situation to a typical higher education where learners' readiness can be determined by the portfolio of courses passed by the learners. We strongly recommend considering a profile of the prospective learners in order to balance learning objectives and learners' proficiency. The self-assessment questionnaires may provide useful information. The key success factor is to ask questions which are relevant to particular skills that will be exercised, e. g., *What tools do you use for detecting*

cyber attacks? instead of *What is your experience with the detection of cyber attacks?*

- *Creating balanced teams* – one of the main aspects of the exercise is to build a sense of teamwork. We advise paying a large amount of attention to creating teams of learners who possess the necessary skills. For instance, if the self-assessment inputs indicate that some learners are experts in one area, it is recommended to distribute them to all teams equally and complement them with experts in another area.
- *Sandbox configuration documents* – continually editing and updating the specification of used systems, network configurations and vulnerabilities is crucial for the successful and smooth preparation of the sandbox. The description should be done using an automation tool such as Ansible [28] to assure its long-term maintainability. Any static documentation (e. g., a wiki page, readme file) is error prone, and becomes outdated very soon.

B. Dry run

The dry run is a complete test of the proposed cyber exercise to get diverse feedback on it. We invite different groups of learners (testing Blue teams) that participate in a pilot exercise. Dry run follows the same schedule and timing as final exercise to rehearse the entire scenario and interaction between Red, White and Green teams, even though it consumes a considerable amount of manpower. It is a mix of *Do* and *Adjust* PDCA phases.

We learned that adjusting the scoring system based on the dry run might be misleading if the expertise and size of the Blue teams participating in the dry run is not similar to learners. The progress of the dry run may be also influenced by various exercise conditions and events that may not happen in the final execution.

C. Execution

The execution phase starts with a familiarization period that enables Blue teams to learn about the exercise infrastructure that has to be defended. The Red team takes no action in this period, so the Blue teams have an opportunity to harden "their" infrastructure. Then the actual exercise starts according to the scenario that is strictly followed by members of the Red and White teams. Once the exercise ends, representatives of the Red, White and Green teams provide a very short assessment of Blue teams' performance during the whole exercise (hot wash-up). This is very desirable since Blue teams can see the final score and can only estimate the content of the exercise scenario.

We identified five challenges related to the execution phase:

- *The level of guidance by organizers* – although creating balanced teams should help to equate learners' proficiency and exercise difficulty, the learners sometimes struggle even though they try their best individually and as a team. We advocate providing some hints by the White team in order to keep the learners in the exercise flow and not to get frustrated because they are stuck at one point. However, the guidance should be provided to all teams equally to preserve fair play.
- *Exercise situational awareness for learners* – the general aim of the exercise is to detect and mitigate cyber attacks. Providing exercise situational awareness for the learners might be contradictory to this aim. We provide only a basic indication of the learners' performance assessed by the White team and Red team by displaying a real-time total score of all teams on a shared scoreboard. This also proved to be an important factor fuelling participants with stress as well as a competitive mood.
- *Exercise situational awareness for organizers* – situational awareness for the White team is very important in the familiarization period where no attacks are conducted against the infrastructure defended by the Blue teams. At the beginning, all systems are intact. Blue teams then reconfigure them to harden them and prepare the infrastructure for attacks by the Red team. The familiarization period is intentionally short so learners are under pressure and they make a number of mistakes. Monitoring the exercise's infrastructure (by the Green team) enables the White team to provide hints for Blue teams in these cases. However, this does not apply in the exercise itself because there may be states that monitoring evaluated as wrong but they were caused by a proper operational decision by a Blue team.
- *Automation of the attacks and injects* – since the exercise scenario is fixed and rigid, Red and White teams may benefit from semi-automated routines that execute the predefined attacks and injects. However, there might be an unexpected situation in which the assistance of a human operator is essential. For instance, the routines expect a file at the default location but the Blue team moved it to another place during the exercise. In addition, we are not aware of any generator of network traffic that can emulate typical Internet users, and that can be easily deployed in the exercise infrastructure.
- *Service access to the exercise's infrastructure* – to recognize an exercise infrastructure failure from scenario progression (e. g., Red team's attack or Blue team's misconfiguration), the Green team needs a service access to all sandbox components. The service access must be clearly defined in the rulebook, no attack will originate from this account, and the Red team does not have access to this account.

D. Evaluation

The exercise life cycle ends with an evaluation. It consists of an assessment of team actions and performance

during the exercise, feedback survey and evaluation (after-action) workshop for the learners, and gathering lessons learned by the organizers.

The most visible part of this phase is the evaluation workshop attended by the Blue teams which lasts about a half day. Other parts of this phase are done by the White and Red teams and require much more time and preparation effort. The White team assesses e-mail communication during the exercise with respect to the non-technical learning objectives (reporting, information sharing, legal). The Red team prepares an overview of its success in attacks against particular teams and best practices related to the attacks used in the exercise. Both teams benefit from data collected by and entered into the scoring application. Furthermore, the Green team stores all collected logs during the exercise of other teams if needed. Feedback provided by the Blue teams in the survey before the evaluation workshop is also incorporated.

All parts of the evaluation (except gathering the lessons learned by the organizers) can be, again, seen as the PDCA *Plan*, *Do* and *Check* phases and the lessons as an input for the *Adjust*.

Through several runs of the exercise, we realized that learning also happens in the evaluation phase. This applies particularly to novices and learners who rated the exercise as difficult. The evaluation workshop shows the exercise scenario and timeline from the perspective of the Red team and White team. It is the only opportunity when the learners can authoritatively learn about attacks used by the Red team. They can discuss their approach in particular situations and phases. Until this point, they were only able to see the results of their experimentation during the exercise without an explanation of *why* something happened. We, therefore, recommend not to underestimate this part of the exercise and deliver analysis and lessons that will have value for the learners. For instance, a handout with best practices for system hardening might be useful in the daily routine of the participants.

E. Repetition

The repetition phase is an instantiation of the exercise sandbox, the execution of the existing exercise scenario for a new group of learners followed by the evaluation. Using the lessons collected in the previous phase, the repetition can be conducted with much less effort and manpower than the first run. It is also possible to skip the dry run phase after one or two repetitions. The repetition includes all phases of PDCA cycle.

V. Conclusions and future work

We have presented a defence exercise deployed in a cyber range and lessons learned from six runs for about 120 adult learners of various expertise, backgrounds and nationalities. The learners have no previous knowledge of the defended infrastructure and the organizers have very limited information about learner's skills and knowledge before the exercise.

We identified a general life cycle of a cyber defence exercise consisting of five phases: *preparation*, *dry run*, *execution*, *evaluation*, and *repetition*. We have described each phase and highlighted important lessons we have learned. Considering these lessons can minimize trial-and-error effort in the design, development, execution and repetition of an exercise.

A. Experience and lessons learned

Finding the best strategy to achieve a cost-effective and sustainable exercise is a very challenging goal. It is a never-ending trade-off between approaching reality and feasibility. Balancing each part of the life cycle allows the creation of a sustainable exercise that can be iteratively improved.

The preparation phase has the decisive influence on final features of the exercise. It is vital to invest many months of manpower into this phase. All systems emulated in the exercise infrastructure must be ready including exercise content, vulnerabilities, and misconfigurations at the beginning of the exercise.

The initial version of the exercise produced in the preparation phase is not sufficient for executing successfully on its own. It must be complemented with a dry run with real learners. In our experience, the dry run verifies not only the story of the exercise but also the ability to use the exercise in repeatable deployments. Poor documentation can cause a lot of problems when making changes in a complex scenario and delay bug fixing and deployment.

Experience from the past exercises highlighted two challenges that we will investigate in our future work: *i)* how to design prerequisite testing, and *ii)* how to provide deeper feedback to the learners immediately after the exercise.

B. Future work

The limited information about prospective learners of an exercise inspired our future research on diagnostic assessment, particularly testing prerequisites for the exercise. Matching learners proficiency and exercise difficulty is a key success factor of the whole exercise. However, the best current practice is announcing the prerequisite skills and knowledge in free form, or acquiring input by self-assessment questionnaires sent out before the exercise. Both proved to be inaccurate. We are investigating methods of gaining objective information using short quizzes, tests and practical tasks related to the learning objectives of the exercise.

The scoring system produces valuable data that may be used either to compare teams mutually, or to show the progress of a team during the exercise. However, so far, the data has been aggregated to a single scoring board consisting of the current or final scores of all teams. We aim to utilize the scoring data to provide better feedback so that the learners can learn from their mistakes.

We plan to present continuous scoring statistics to the learners immediately after the exercise in a well-considered interactive way and analyse their physical behaviour (e. g., eye-tracking, mouse event recording) in order to catch the interest of the learners. These techniques would expose how much feedback helps them to get insight into the passed exercise. We believe that the improved feedback from the exercise may increase learners' motivation to attend further exercises.

Acknowledgements

This research was supported by the Security Research Programme of the Czech Republic 2015-2020 (BV III/1 – VS) granted by the Ministry of the Interior of the Czech Republic under No. VI20162019014 – Simulation, detection, and mitigation of cyber threats endangering critical infrastructure.

Access to the CERIT-SC computing and storage facilities supported from European Regional Development Fund-Project "CERIT Scientific Cloud" (No. CZ.02.1.01/0.0/0.0/16_013/0001802), is greatly appreciated.

The Cyber Czech exercise series was designed, developed and carried out in cooperation with the National Cyber Security Centre (NCSC), a part of the National Security Authority of the Czech Republic.

References

[1] A. Pras, A. Sperotto, G. Moura, I. Drago, R. Barbosa, R. Sadre, R. Schmidt, and R. Hofstede, "Attacks by "anonymous" wikileaks proponents not anonymous," University of Twente, Centre for Telematics and Information Technology (CTIT), Tech. Rep., 2010.

[2] M. Corp., "Exposing one of China's cyber espionage units – Mandiant APT1 report," Tech. Rep., 2013. [Online]. Available: http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf

[3] Burning Glass Tech, "Job market intelligence: Cybersecurity jobs," Tech. Rep., 2015. [Online]. Available: http://burning-glass.com/wp-content/uploads/Cybersecurity_Jobs_Report_2015.pdf

[4] Cisco Systems, "Cisco 2014 annual security report," Tech. Rep., 2014. [Online]. Available: http://www.cisco.com/web/offer/gist_ty2_asset/Cisco_2014_ASR.pdf

[5] A. Davis, T. Leek, M. Zhivich, K. Gwinnup, and W. Leonard, "The fun and future of ctf," in *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*. San Diego, CA: USENIX Association, 2014.

[6] NATO Cooperative Cyber Defence Centre of Excellence, "Cyber defence exercises." [Online]. Available: http://ccdcoe.org/event/cyber-defence-exercises.html

[7] J. Mirkovic and P. A. H. Peterson, "Class capture-the-flag exercises," in *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*. San Diego, CA: USENIX Association, 2014.

[10] "Societal security – guidelines for exercises," International Organization for Standardization, Geneva, CH, Standard, Sep. 2013.

[8] M. Granåsen and D. Andersson, "Measuring team effectiveness in cyber-defense exercises: a cross-disciplinary case study," *Cognition, Technology & Work*, vol. 18, no. 1, pp. 121–143, 2016.

[9] E. G. Díez, D. F. Pereira, M. A. L. Merino, H. R. Suárez, and D. B. Juan, "Cyber exercises taxonomy," INCIBE, Tech. Rep. [Online]. Available: https://www.incibe.es/extfrontinteco/img/File/intecocert/EstudiosInformes/incibe_cyberexercises_taxonomy.pdf

[11] B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar, "An integrated experimental environment for distributed systems and networks," Boston, MA, Dec. 2002, pp. 255–270.

[12] "Cyber virtual ad hoc network (CyberVAN)." [Online]. Available: http://www.appcomsci.com/research/tools/cybervan

[13] L. Buttyán, M. Félegyházi, and G. Pék, "Mentoring talent in IT security–A case study," in *2016 USENIX Workshop on Advances in Security Education (ASE 16)*, 2016.

[14] "Avatao." [Online]. Available: https://avatao.com

[15] CTF365, "Capture the flag 365." [Online]. Available: https://ctf365.com

[16] Security Competence, "Hacking-lab." [Online]. Available: http://www.hacking-lab-ctf.com/technical.html

[17] G. Vigna, K. Borgolte, J. Corbetta, A. Doupe, Y. Fratantonio, L. Invernizzi, D. Kirat, and Y. Shoshitaishvili, "Ten years of iCTF: The good, the bad, and the ugly," in *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*, 2014.

[18] A. S. Raj, B. Alangot, S. Prabhu, and K. Achuthan, "Scalable and lightweight ctf infrastructures using application containers," in *2016 USENIX Workshop on Advances in Security Education (ASE 16)*. Austin, TX: USENIX Association, Aug. 2016.

[19] J. Davis and S. Magrath, "A survey of cyber ranges and testbeds," DTIC Document, Tech. Rep., 2013.

[20] J. Mirkovic, T. V. Benzel, T. Faber, R. Braden, J. T. Wroclawski, and S. Schwab, "The Deter Project," 2010.

[21] T. Benzel, "The science of cyber security experimentation: The DETER project," in *Proceedings of the 27th Annual Computer Security Applications Conference*. ACM, 2011, pp. 137–148.

[22] B. Ferguson, A. Tall, and D. Olsen, "National cyber range overview," in *2014 IEEE Military Communications Conference*, Oct 2014, pp. 123–128.

[23] MCR, "The michigan cyber range." [Online]. Available: https://www.merit.edu/cyberrange/

[24] L. Rossey, "SimSpace cyber range," aCSAC 2015 Panel: Cyber Experimentation of the Future (CEF): Catalyzing a New Generation of Experimental Cybersecurity Research. [Online]. Available: https://www.acsac.org/2015/program/ACSAC%202015%20CEF%20Panel%20-%20Rossey.pdf

[25] "EDURange." [Online]. Available: http://www.edurange.org

[26] J. Vykopal, R. Ošlejšek, P. Čeleda, M. Vizváry, and D. Tovarňák, "KYPO Cyber Range: Design and Use Cases," in *International Conference on Software Technologies (ICSOFT'17)*, Madrid, Spain, 2017.

[27] NATO cooperative cyber defence centre of excellence, "Locked shields." [Online]. Available: http://ccdcoe.org/event/cyber-defence-exercises.html

[28] Red Hat, "Ansible." [Online]. Available: https://www.ansible.com