Journal Pre-proof

Cyber Ranges and Security Testbeds: Scenarios, Functions, Tools and Architecture

Muhammad Mudassar Yamin, Basel Katt, Vasileios Gkioulos

PII: S0167-4048(19)30180-4
DOI: https://doi.org/10.1016/j.cose.2019.101636
Reference: COSE 101636

To appear in: *Computers & Security*

Received date: 21 March 2019
Revised date: 8 July 2019
Accepted date: 6 October 2019

Please cite this article as: Muhammad Mudassar Yamin, Basel Katt, Vasileios Gkioulos, Cyber Ranges and Security Testbeds: Scenarios, Functions, Tools and Architecture, *Computers & Security* (2019), doi: https://doi.org/10.1016/j.cose.2019.101636

Cyber Ranges and Security Testbeds: Scenarios, Functions, Tools and Architecture

Muhammad Mudassar Yamin*, Basel Katt, and Vasileios Gkioulos
Department of Information Security,
Teknologivegen 22, 2815 Gjøvik 2815 Gjøvik,
Oppland, Norway.
* Corresponding author. muhammad.m.yamin@ntnu.no

Preprint submitted to Elsevier *October 7, 2019*

Cyber Ranges and Security Testbeds: Scenarios, Functions, Tools and Architecture

Muhammad Mudassar Yamin, Basel Katt, and Vasileios Gkioulos

March 2018

Abstract

The first line of defense against cyber threats and cyber crimes is to be aware and get ready, e.g., through cyber security training. Training can have two forms, the first is directed towards security professionals and aims at improving understanding of the latest threats and increasing skill levels in defending and mitigating against them. The second form of training, which used to attract less attention, aims at increasing cyber security awareness among non-security professionals and the general public. Conducting such training programs requires dedicated testbeds and infrastructures that help realizing and executing the training scenarios and provide a playground for the trainees. A *cyber range* is an environment that aims at providing such testbeds. The purpose of this paper is to study the concept of a cyber range, and provide a systematic literature review that covers unclassified cyber ranges and security testbeds. In this study we develop a taxonomy for cyber range systems and evaluate the current literature focusing on architecture and scenarios, but including also capabilities, roles, tools and evaluation criteria. The results of this study can be used as a baseline for future initiatives towards the development and evaluation of cyber ranges in accordance with existing best practices and lessons learned from contemporary research and developments.

1 Introduction

The recent security incidents worldwide have shown that there is an increase in the complexity and severity of cyber security threats. The attackers become more organized and the attack vectors are using more advanced and automated techniques and tools. The first line of defense against such attacks is increasing cyber security awareness in the public and security skills among the security professionals, in order to be ready and aware of the latest threat techniques and tools. These training programs include the execution of cyber security labs and exercises. In general terms, we define a cyber security exercise as a training exercise that runs attack and/or defense scenarios on virtual and/or physical environments with the aim of improving the attack and/or defence understandings and skills of the participants. Different groups of people are

involved in preparing and executing such exercises. A groups of individuals, known as a white team, creates the training environment. Another group, known as a red team, tries to exploit vulnerabilities present in the environment, while a third group, known as a blue team, tries to defend the environment and prevent attacks. These are the main basic roles for those who are involved in an exercise. More comprehensive list of all roles within an exercises is discussed later in the chapter. Please note that we use the term *security exercise* for any practical training or awareness activity.

Researchers divided a security exercise life cycle in five phases [113], which are *preparation*, *dry run*, *execution*, *evaluation*, and *repetition*. In the first phase the exercise objectives, scenario story, scoring method, and the environment will be set up. In the dry run phase, the developed environment will be tested according to the exercise objectives. The execution phase involve running the exercise, in which the participants in the attacking and/or defending side will try to achieve their objectives. In the evaluation phase, the performance of the participants will be assessed based on the scoring method and learning objectives. Finally, in the last phase, the environment is cleaned and the whole process is repeated for a new exercise. It has been observed [113] that security exercises are usually conducted and evaluated (execution and evaluation phases) in few hours up to a few days, while the preparation and dry run often take up to months for completion. This makes security exercises very costly and time consuming to be used in large scale to help reducing the growing cyber security skills gap [40].

In order to maintain and manage security exercises and their environment, a cyber range concept has been proposed. Recently, the concept and the term has attracted a great attention, but has been used differently in different contexts. Some use it to refer to a virtual environment, and others include other physical elements to a cyber range. It can refer to a university lab environment, or it can refer to a classified security exercise environment. There has been some attempts to study and classify the concept of a cyber range, e.g., the survey conducted by the Australian defense in 2013 [28]. Such studies provide a general background and classification of the term, though, (1) they do not cover all aspects of a cyber range system, e.g., architecture, management or scenarios, (2) they are outdated when it comes to cyber range technologies and tools, and (3) they do not discuss research trends and directions. Others, like [53] and [83] are not generic enough and focus on specific exercise domains, like smart grids. To cover the gap in the literature, we conducted a systematic literature review on the topic of cyber range systems. The goals is to analyze the current state of the art within the topic of unclassified cyber ranges and security testbeds, and make recommendations regarding the architecture, capabilities, tools, the testing and training process, scenarios, and evaluation. The result can be used as a baseline for future initiatives towards the development and evaluation of cyber ranges in accordance with existing best practices and lessons learned from contemporary implementations.

The rest of this paper is structured as follows. In this next section, we present the related work covering the similar surveys and reviews conducted

on this topics. In section 3, we present the methodology and in section 4 we discuss the results. In section 5 we synthesize the result and present a general purpose architecture for a cyber range and summarize the research trends and directions. Finally, in section 6 we discuss and conclude the paper.

2 Related Work

During planning and writing this article, no other systematic literature review was found by the authors on the topic of cyber ranges and security testbeds. Yet, a multitude of survey articles has been identified with focus on specific application domains such as industrial control systems, mobile ad-hoc networks and cyber physical systems. Leblanc et al. [63] in 2011 presented an overview of cyber attack and computer network operations simulation and modeling approaches. The discussed approaches have been identified within the open literature, and originate from governmental and academic efforts as well as from the private sector. These include, but are not limited to, ARENA, RINSE (Real-Time Immersive Network Simulation Environment), SECUSIM, and NetENGINE. In respect to research activities driven by the private sector and academia, the authors found that there are substantial efforts focused on cyber attack modeling, with constructive automated simulations. The results enabled the discovery of cyber attack patterns, with accuracy that is primarily dependent on the utilized models. Yet, the authors noticed that the governing parameters for most of these models are not validated against real world scenarios. Therefore, they mostly focused on specific artificial educational scenarios, rather than analysis of realistic cyber attacks in general. Furthermore, they overlooked also cascading effects on organizational or national scale.

Siaterlis et al. [94] in 2009 investigated available software for the creation of testbeds for Internet security research. The authors identified that numerous publications refer to prototypes rather than to software that is ready to be used for the creation of testbeds. Accordingly, they proposed a framework for feature based evaluation of the available software, as well as, they provided a literature review and comparison of state-of-the-art tools. This study excluded platforms that (i) share computational resources, (ii) focus only on simulation, (iii) are specific to wireless or sensor networks, (iv) run on a single computer, and (v) use custom hardware. The proposed framework consists of 13 basic and 6 compound features, including (i) distinction of roles, (ii) remote access, (iii) virtualization, and (iv) clean reconfiguration. The authors categorized their findings to overlay testbeds, including Planetlab and X-Bone, and cluster testbeds, including Grid'5000, Emulab, and ModelNet. They concluded that Emulab and Planetlab provide the most mature solutions for each testbed type and sufficient documentation for the development of dedicated testbeds, while Flexlab seeks to combine the best characteristics of the two approaches.

Davis et al. [28] provided a survey of unclassified cyber ranges and testbeds, in a study completed in October 2013. The article provides an overview of background information in terms of supported functionalities and terminology,

and also covers specific implementations originating from the military, public governments, and academia. SECUSIM, RINSE, ARENA, and LARIAT are some of the testbeds covered. The authors promoted hardware emulation as the most realistic approach, with simulations, on the other hand, providing increased flexibility and scalability advantages. Yet, as the study suggests, the middle ground providing parameterized support for emulation, simulation, and virtualization is increasingly explored, highlighting again Emulab and DETER as the most mature solutions.

Holm et al. [53], Sun et al. [107], Qassim et al. [83], and Cintuglu et al. [26] focused on testbeds dedicated to cyber physical systems, such as industrial control systems, SCADA, and the power grid. The articles investigate testbeds that have been proposed for scientific research and educational activities in aspects related to objectives, capabilities, architectural designs and integrated components, as well as implementation techniques for satisfying requirements. The authors also referred to these articles with explicit design and integration recommendations. Specifically, although the examined testbeds seem to target objectives such as vulnerability analysis, education, and tests of defensive mechanisms, these are not thoroughly described. In order for them to relate to specific architectural decisions, they must be refined and aligned with specific target vulnerabilities.

Balenson et al. [9] focused on cyber security experimentation for the future. They worked on devising fundamental and new experimentation techniques for cyber security research. They concluded that new methods of research is required in cyber security focusing on just hardware and software is not enough. A community driven approach is required to constantly train the workforce in a dynamic cyber security enviroment. Carnegie Mellon University has developed a LMS (Learning Management System) which is called *StepForward* [18]. It provides the opportunity to teach students both theoretical and practical cyber security skill set in a realistic environment by combining multiple choice questions with emulated labs. In term of cyber security competitions that use different cyber ranges and security testebeds, a comprehensive list is maintained at *cybersecuritydegrees* [1]. Cyber security competitions are a good way to measure the effectiveness of cyber security training.

3 Methodology

The systematic literature review is a research review that aims at identifying, evaluating and synthesizing the existing literature of scientific work regarding a particular research question or topic. We decided to follow this method because it results in a credible, objective and unbiased evaluation of the current literature. This study has been conducted in accordance with the protocol described by Okoli et al [24] in their article *"A Guide to Conducting a Systematic Literature Review of Information Systems Research"*. The protocol consists of eight consecutive steps, namely: (1) Define the purpose of the literature review, (2) establish a protocol among the participants, (3) search the literature, (4) per-

form practical literature screening, (5) perform quality appraisal, (6) perform data extraction, (7) synthesize the results, and (8) write the review. Three researchers participated in the literature review. In the following paragraphs, we provide the required insights of the adopted methodology, in order to enhance the readability of the following sections and support future derivative or continuation studies.

3.1 Purpose of the literature review

The main purpose of this literature review is to study the concept of a cyber range system. Various aspects of a cyber range will be considered and a taxonomy will be created. Specifically, the objectives of this systematic literature review can be summarized as follows:

1. To identify and classify the capabilities and functionalities deployed within contemporary cyber ranges and security testbeds.
2. To collect and critically evaluate existing cyber ranges and security testbeds' architectural models.
3. To identify and classify scenarios, for training or testing, applied in cyber ranges and security testbeds.
4. To identify the different roles and teams associated with the execution of an exercise in a cyber range.
5. To identify and classify hardware and software tools utilized within contemporary cyber ranges and security testbeds.
6. To identify methods to evaluate different cyber ranges against a standard.
7. To study the research trends and directions on the topic of cyber ranges and security testbeds.

3.2 Establishing the review protocol

Three researchers participated in this systematic literature review from the period between March 2018 until January 2019. At the beginning a discussion round resulted in the selection of the concrete methodology. The methodology was shared and studied by all members. After the selection and the study of the methodology, a concrete protocol for the execution of the review was established and a cloud based repository was created to maintain temporary files and document the conducted steps. Templates for documentations, data extraction, and storing the results according to the established protocol were created as well.

3.3 Searching the literature

We followed the established protocol for systematic literature review in order to help the reproducibility of the study [24] and provided the details in comprehensive methodology. We employed keywords based search technique in order to identify relevant literature. The keywords were selected very carefully in order to fulfill the purpose of the review described in 3.1. We performed a preliminary search using only the term "cyber range" and the results were not comprehensive. We noticed that there are some work that uses the name security testbed and security exercise when talking about a "cyber range" system. So, we decided to use the words "testbed" and "exercise". The collection of the literature was undertaken in accordance with the following parameters:

- Examined scientific databases: ACM digital library, IEEE Xplore, ScienceDirect, Springer Link, and Wiley online library.
- Utilized keywords (advanced search): "Cyber Range", "Security"+"Testbed", "Security"+"Test-bed", "Security Exercise".
- Publication period: 15 years (2002 - 2018).
- The total period of the literature review: March 2018 - January 2019.

3.3.1 Search criteria

The search for security testbed results in a large amount of work, in which researchers conducted an experiment and they used a specific testbed for that purpose. These works were not of an interest for this review, and accordingly, we developed the list of rigorous inclusion and exclusion criteria. Thus, we listed the topics in which security testbeds were only mentioned to describe an experiment that was conducted in a particular domain, e.g., robots, UAV, and RFID testbeds. The application domains that can be included in the survey are vast, ranging from chemical-focused laboratories, to environmental systems. Covering all possible domains in one survey is not feasible and not possible. Therefore, we had to exclude some of the application domains to make it feasible, taking the maturity of the domain and the security relevance as two factors in this decision. Based on an internal discussion among the researchers, we decided on the list of inclusion and exclusion criteria that cover most important domains (not all), but make the survey feasible. For example, we cover the smart grid and industrial SCADA systems, but at the same time, we excluded transport systems, UAV, and robotics. The same applies for mobile infrastructure. In this case, we focused on application layer in the mobile testbeds, e.g., BYOD testbed scenarios, but we excluded infrastructure focused testbeds, like 4G/5G/GSM, and WIMAX testbeds. Thus, the identified literature was based on the following inclusion and exclusion criteria:

1. Inclusion criteria: The following inclusion criteria were applied in the review.

- Articles written in English.
- Security relevant testbed and exercises. Either presenting a whole cyber-range or a section/component of a cyber-range.
- IoT (Internet of Things) related testbeds.
- CPS (Cyber Physical Systems) and SCADA related testbeds.
- Articles related to cyber-range federation.
- Articles related to mobile applications testbeds.

2. Exclusion criteria: Based on the aforementioned discussion, in the following is the list of criteria we develop to filter out papers that are not within the scope of this review.

   - Articles that mention testbeds in the context of other work. The focus must be on the testbed.
   - Testbeds for UAV (Unmanned Aerial Vehicle).
   - Testbeds for RFID, NFC, and WIMAX.
   - Testbeds for cryptographic protocols.
   - Testbeds for robots .
   - Testbeds for trust related issues.
   - Testbeds focusing on security of structures, transportation, and security/safety of persons.
   - Testbeds focusing on climate change and the environment.
   - Testbeds for simulation of underwater sensor.
   - Conference abstracts, book reviews, conference info, discussion, editorials, mini reviews, news, short communications.

3. Quality appraisal: The focus of this paper is to study cyber ranges and security testbeds as a whole, in order to give insights to those who are designing, building, researching or operating a cyber ranges and security testbeds. For this reason, a relevant quality appraisal criteria is defined is to cover and study the cyber ranges and security testbeds as whole. This survey can be followed by other surveys that focuses on a particular aspect of a cyber ranges and security testbeds like scenarios, teaming, scoring etc.

   To ensure significant and quality contributions, we established an additional filtering step. We decided on the following list of topics related to general cyber range investigation, which are part of the taxonomy that we propose later in the paper. We noticed in an initial screening, that papers that use testbeds in the context of another research that is not related to the testbed itself, mentioned the scenario and an additional aspect, like scoring, monitoring, or management, depending on the research conducted. This means that papers that mentioned only one or two of the topics we specify, are not relevant. Therefore, significance and relevance were decided if articles include in their investigation at least three of the following five areas or topics of investigation:

(a) Scenarios (architecture and story/behavior)

(b) Monitoring and logging

(c) Teaming

(d) Scoring

(e) Management (Id management, resource management, cyber range management, life cycle management)

Additionally, the following quality assurance criteria were taken into consideration.

(a) Originality of the work.

(b) Quality of presentation.

(c) Scientific soundness and method.

(d) Papers that have been cited should be included in the survey. This rule is exempted from papers that were published recent, i.e., less or equal then two years. The citation data as of August 10th 2018 is parented in appendix table 13.

3.4 Practical literature screening

Based on the aforementioned steps and criteria, we conducted the practical literature screening. The following rounds were resulted.

1. Round 1: Collection of the literature was conducted in March 30th. It resulted in a total entries of **385**.
2. Round 2: Elimination of duplicates was conducted in April 25th, and resulted in a total entries of **310**.
3. Round 3: Back tracing additional entries from the citations of the current articles was conducted in June 20th. It resulted in a total number of entries **341**.
4. Round 4: Quality appraisal was conducted on August 10th, and resulted in the total number of articles **100**.

3.5 Classification and data extraction

Based on the work we have done in developing a cyber range and after the first screening of the literature, we propose an initial taxonomy to classify cyber ranges as shown in figure 1. A new updated taxonomy is developed after the survey was conducted and will be presented in section 4.2. In the following is a short description of each concept.

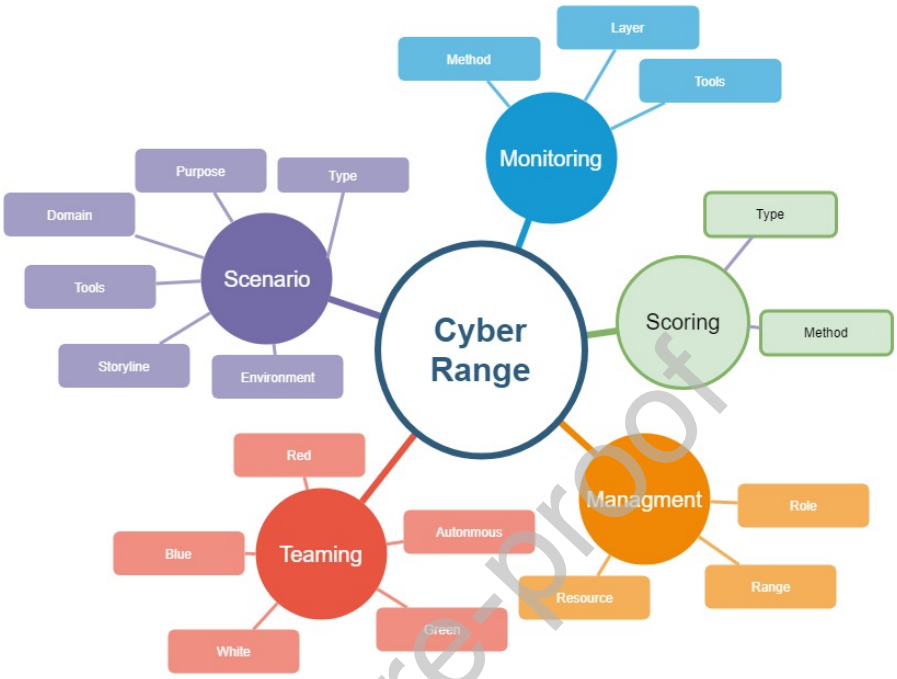

Figure 1: Cyber Range taxonomy

1. **Scenarios**
   A scenario defines the execution environment as well as the storyline that indicates the execution steps of a test or a training exercise. It accurately represents the operational environment and training requirements, and drives training execution to ensure the achievement of training objectives. The scenario describes and provides documentation, summaries, action orders, etc., to ensure the representative operational context supports testing and training objectives [104]. We classify a scenario to extract information about what is the purpose of the exercise, or test? Where an exercise, or a test, is executed? How an exercise, or a test, is executed? And which tools are used in the execution of a scenario? answers to these questions are given below.

   (a) **Purpose**
   The purpose explains what are the objectives of the scenario, i.e. the execution of a cyber security training exercise or the experimentation validation of new cyber security tools and techniques. Based upon the scenario objectives, scenario environment is developed, details of which are given below:

   (b) **Environment**
   The scenario environment is the topology where the scenario is exe-

cuted. The scenario depends upon the exercise and experiment objectives. If the exercise is an operation-based, then the environment will be a technical infrastructure, i.e., computer based, physical, virtualized or hybrid. If the exercise is a table-top or discussion based the environment can be non computer based [48]. In a table top based cyber security exercise a cyber scenario is discussed and the decision making ability of the exercise participants is evaluated. It can be computer aided or can be executed without the use of any digital equipment.

(c) **Storyline**
A storyline of a scenario tells a single or multiple stories about how the exercise will be executed. It includes the development of relevant actions and events that constitute the scenario and how these are connected to generate the whole narrative of a scenario. This allows the overall understating and controlling of a big technical scenario, and gives the ability to critically evaluate the exercise, OR test, outcome [104]. In term of experimental validation of new technologies, single or multiple test case can be executed for research are investigation.

(d) **Type**
The type of the scenario indicates whether the scenario is static or dynamic. We define a scenario to be static, if it includes a static environment, and no changes are applied during the execution of the exercise. This means that the storyline does not include any dynamic components that changes over time. A dynamic scenarios are scenarios that include, besides the static environment, a dynamic component that will make changes during the execution of the scenario. For example, a simulator, or a traffic generator that can be injected, or executed, during the exercise.

(e) **Domain**
The domain indicates the application domain of the scenario, e.g., IoT, network, cloud etc.

(f) **Tools**
The tools which are used in the development of scenario. This includes the tools which are needed for the creation of the environment of the scenario, or the tools which are used in the development of a storyline.

2. **Monitoring**
Monitoring includes the methods, the tools and the layers at which real time monitoring of cyber security exercises and tests are performed [104]. Monitoring of cyber security exercise participants is performed by designated observers [57]. The methods that the observers employ, the tools that they use and the layers at which they perform monitoring are further classified:

(a) **Methods**
This classifies methods employed to monitor the cyber security exercise and tests, i.e., how the cyber security exercise, or the test, is monitored. Either automatically with the use of tools that gather data for analysis, or manually by human observers.

(b) **Tools**
This classifies the software and hardware tools that can be used for monitoring of cyber security exercises and tests. The software and hardware tools may include security information and event management (SIEM) solutions and intrusion detection systems etc.

(c) **Layers**
This classifies the layer at which monitoring is being performed. Depending on the type of an exercise, monitoring can be performed at multiple TCP/IP layers, in case of an operation-based exercise; or at an abstract social layer, in case of a table-top exercise.

3. **Teaming**
In a cyber security exercise teaming includes individual and group of individual that design, develop, manage and participate in a cyber security exercise or a test [91]. Based upon a team's role in a cyber security exercise different colors are assigned to them to identify their role [114]. Details of which are given below:

(a) **Red team**
Red teaming is a form of information security assessment in which cyber- security adversaries are modeled to identify vulnerability present in a system during an exercise or a test [119]. The red team is responsible to identify and exploit potential vulnerabilities that are present in the exercise environment.

(b) **Blue team**
Blue teaming is a form of active defense against an active attack on a cyber security exercise and test environment [116]. The blue team is responsible to identify and patch potential vulnerabilities that can be exploited by a red team.

(c) **White team**
A white team designs the exercise and experiment scenario, objectives, rules and evaluation criteria. They set a set of rules of engagement between red and blue team, inject the vulnerabilities in the environment for patching and exploitation; and sometimes they act as instructors to give hints to the participating teams [114].

(d) **Green team**
A green team is responsible for the development, monitoring and maintenance of the exercise infrastructure designed by the white team. They are also responsible for fixing bugs and crashes in the infrastructure occurred during an exercise execution [114].

(e) **Autonomous teams**
Team roles that are being automated by different tools and techniques are considered as autonomous teams. For example Secgen [92] is used for the automation of scenario environment development which is the role of green team and SVED [52] is used for the role automation of a red team.

In some cyber security exercises additional teams are included, which are exercises specific and not present in cyber security exercise life cycle [114]. Details of which are given below:

(a) **Orange Team**
Orange team members assign different technical tasks to blue team members during the exercise. Blue team members can earn points if they are able to successfully complete the tasks.

(b) **Purple Team**
Purple teams perform the communication role between multiple exercises teams. They do information sharing to increase the exercise effectiveness. This enhances the effectiveness of a red team in attacking the exercises environment and increases the capability a blue team in defending the network.

(c) **Yellow Team**
Yellow team members simulates the behavior of normal users that are using the infrastructure created by the green team. They performs the tasks like generating legitimate network traffic which can be used by red and blue teams in attack and defense.

4. **Scoring**
Scoring uses data from monitoring systems in order to give performance related semantics to the low level technical events observed during monitoring of cyber security exercises and tests. Some scoring indicators might not depend on technical monitoring events, like flags or over-the shoulder evaluation mechanisms. The scoring mechanism is also used to measure the teams and test progress during an exercise, or a test [114]. The methods and tools used in the scoring mechanism are further classified:

(a) **Methods**
This classifies whether the scoring is done based upon achieving a specific objective, i.e flags, or it is done by analyzing logs that are generated during cyber security exercises or tests.

(b) **Tools**
This classifies the software and hardware tools that are used for scoring of cyber security exercises or tests. The tools may include flags submission dashboards, log analyzers, etc.

5. **Management**
Management involves the assignment of roles and duties to individuals

and teams. Allocation of computational and other resources required for conducting a cyber security exercise, or a test, and the overall management of the cyber range.

(a) **Role management**
Role management classifies the methods, tools and techniques with which the identities and roles of individuals and teams involved in a cyber security exercise, or a test, are managed.

(b) **Resource management**
Resource management classifies the computational resources like processing frequency, memory and disk space required for conducting cyber security exercise, or a test.

(c) **Range management**
Range management classifies the methods, tools and techniques with which the holistic view of overall cyber security exercise, or a test, is presented in portals and dashboards.

4 Analysis of Results

In this section we present and discuss the results of the literature review. First, we discuss how the main capabilities identified in the taxonomy presented in section 3.5 have been investigated, or considered in the literature. Then, we discuss, in more details, the architecture of the contemporary cyber ranges, scenarios, teaming, evaluation criteria, tools used, and future directions presented in major work.

4.1 General capabilities

As per our selection strategy presented in 3.3, a classification of the capabilities and functionalities deployed within contemporary cyber ranges and security testbeds is presented in figure 2 and table 1. We identified that the capability that was mostly investigated in the literature is *scenarios* with 94 papers that include details about scenarios. The second most prominent capability is management with 91 papers. Then there were 86 papers that have details about the monitoring infrastructure, 41 papers contain details about teams, and only 26 papers have details about the scoring mechanism.

Paper	Scenarios	Monitoring	Teaming	Scoring	Mng.
[23][114][78][113][19][34][70] [28][64]	✓	✓	✓	✓	✓

[15][20][29][107][35][37] [56][21] [86] [5] [39][33][120] [74][42][69][108][93][47][88] [8][3][4][67][7][84][106] [112][31][75][65][81][103][95] [41][49][13][109][73][58][55][50] [17][98][44][38][115][25] [59][26][71] [54][10]	✓	✓			✓
[68][110][118][117][80] [32][12][27][11][76]	✓	✓	✓		✓
[36][82][72][66][16][77][121]	✓		✓		✓
[105][96]	✓	✓		✓	✓
[60]	✓	✓		✓	
[46][97]	✓	✓	✓		
[45][22]		✓	✓		✓
[100][87]	✓			✓	✓
[101]		✓	✓	✓	✓
[51][90][102]	✓		✓	✓	✓
[99][6][61]	✓	✓	✓	✓	
[30][85]	✓		✓	✓	
[111]		✓	✓	✓	
[2]		✓		✓	✓
[89]		✓	✓	✓	✓

Table 1: Capabilities and functionalities deployed with in contemporary cyber ranges and security testbeds

In order to analyze the evolution of these different capabilities over time, figure 2 depicts how the interest of different capabilities has increased steadily, with few exceptions, since 2002. It can be noticed that in the period between 2007 and 2008 the number of publications dropped, and then continued increasing in 2009 until 2017. This is correlated with the fact the major cyber ranges, like the US National Cyber Range have started development in the period between 2008 and 2009. Before that date, most of the work was conducted in terms of general purpose security testbeds. Around the time the US National Cyber Range [79], among FIRE(Future Internet Research and Experimentation) [43] in Europe started which aimed to interconnect existing security testbeds. Due to which many researchers started looking at the new "Cyber Range" concept, which explains the dip in publication around 2008. It is worth mentioning that due to the fact that the screening happened in the second quarter in 2018, the figures related to 2018 is not complete. Also, there were few papers that were found during the search with publication date scheduled in 2019.

4.2 New Taxonomy

The taxonomy presented in section 3.5 is good for identifying the general capabilities of cyber ranges and security testbeds. However, After reviewing the se-

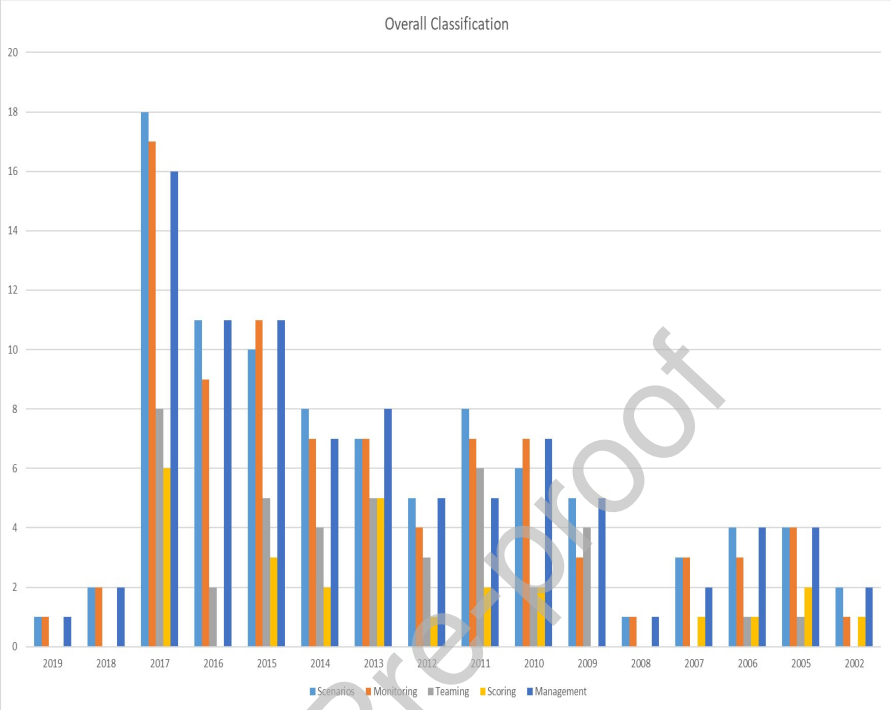

Figure 2: Overall classification of cyber ranges and security testbeds capabilities with respect to years

lected papers and analyzing the collected data, we identified that the taxonomy that we used to identify the general capabilities was not sufficient in presenting cyber ranges and security testbeds functionalities in depth. Therefore, we are proposing a new updated taxonomy for presenting the functionality of cyber ranges and security testbeds based upon the collected data. The developed taxonomy is parented in figure 3. In this section we will focus our discussion on the new elements that were added to the new taxonomy. We will refer to the papers that included information about these new concepts. In general it is worth mentioning the following two main changes compared to the initial taxonomy. First, due to its importance and being related to different other concepts, environment is presented on its own, separately from scenarios. Second, we added the learning concept, as we noticed that learning modules were mentioned repeatedly in cyber ranges. Scoring is considered as a sub-element of the learning module, and thus added as a sub-concept to the learning concept. Apart from that, we expanded the scenario concept with the scenario lifecycle, and the management with command&control and data storage concepts.

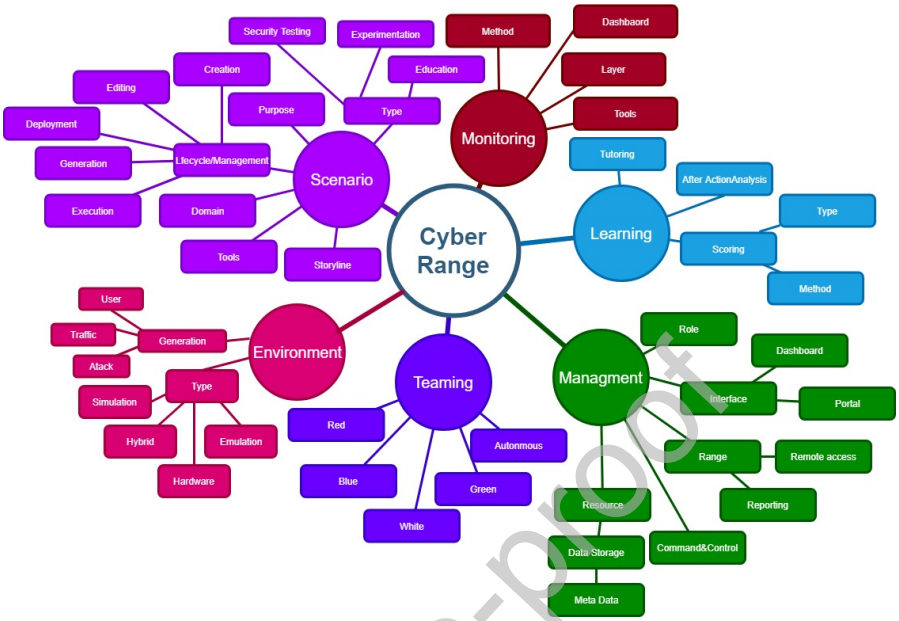

Figure 3: Updated taxonomy of a cyber range

4.2.1 Scenarios

In this section, first, we discuss the cyber security scenario lifecycle management. It involves creating, generating, editing, deploying and executing a cyber security scenario.The following work [78, 72, 108, 39, 121, 45, 54, 4] have specialized components in their architecture to create and edit cyber security scenarios. They mostly have a designer dashboard in which different components of a scenario are presented, and can be used to develop new scenarios. The works in [115, 65] have components to generate cyber security scenarios using different automation techniques. The scenarios are created mostly in a human and machine readable language like XML and JSON, which is then executed on a compiler to deploy the scenario. The works presented in the these papers [111, 39, 50, 54, 108, 72] included special scenario deployment component which is responsible for deploying network resources, like routers and firewalls, and relevant applications, like vulnerable software. For scenario execution, [42, 88, 4, 75, 34] have module that can control the scenario flow, like start, stop and pause scenario execution. Works in [118, 4] have orchestration modules that combine multiple components to execute a scenario. Finally, [42, 4, 56, 61, 10, 39, 34, 90, 98, 21] have components that are used to generate different events within the scenario execution to make the scenario more dynamic and realistic. These events can be the launch of automatic attacks, like in [34, 39, 21], or can be represent traffic generation, like in [90, 61].

Figure 4 shows the evolution of the different purposes of scenarios, i.e., test-

ing, education, and experiment. It can be seen that testing and education are gaining a lot of attention in the last few years, particularly testing. With respect

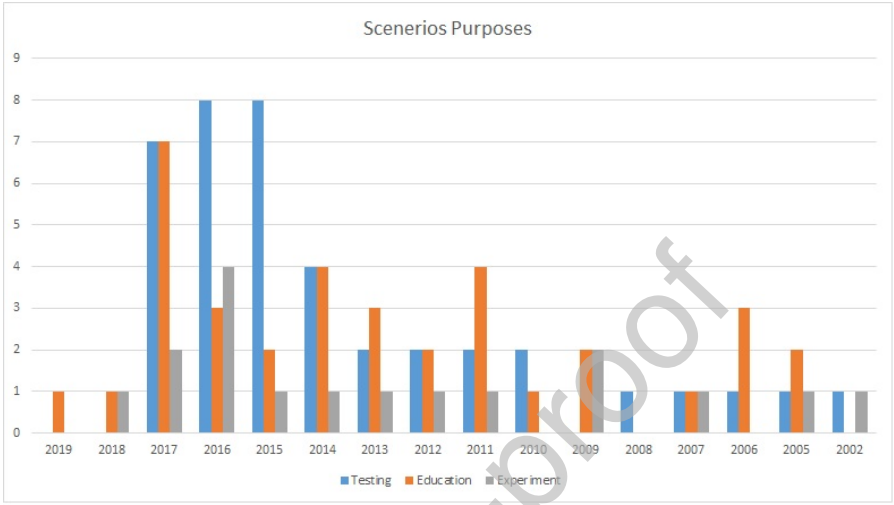

Figure 4: Classification of cyber-ranges and security testbeds based upon the scenarios purpose

to the scenario type, we can distinguish between both static and dynamic scenarios (cf. section 3). Figure 5 shows the evolution of scenario types discussed in the reviewed papers. It can be seen that before 2011 static scenarios, in which the scenario story was not discussed but included only the static topology, was dominant. Since 2011, cyber range scenarios started to add the dynamic component, in which the storyline and the behavior are specified. This shows an advancement in the specification and execution of scenarios in cyber ranges and security testbeds.

Finally, when it comes to the domains of the scenarios, figure 6 shows the different application domains, in which scenarios are specified. Those domains are (1) hybrid network applications, (2) Networking, (3) SCADA systems, (4) social engineering, (5) IoT systems, (6) critical infrastructure, (7) Cloud based systems, and (8) autonomous systems. The figure indicates that networking systems were the main application domain for cyber ranges and security testbeds, SCADA system started to gain attention from 2010, and in recent year cyber ranges and security testbeds have covered most application domain aforementioned. In table 2 we present scenario samples from each application domain, including the purpose, the environment, the storyline topic, and tools used.

Id	Domain	Paper	Purpose	Environment	Storyline	Tools

1	Hybrid Network and Application	[50]	Education	Hybrid	Network topology configuration for students	XEN, CISCO routers
2	Networks	[12]	Experiment	Emulation	DDoS, Worm Behavior, Early Routing Security experiments	Emulab
3	IOT	[98]	Testing	Hardware	Bring your own device scenario testing for enterprises	Smart Wacthes, google glass, printers
4	Critical Infrastructure	[44]	Testing	Emulation	DoS attack on a powergrid	Emulab
5	SCADA	[38]	Experiment	Hardware	DoS, ICT worm, Phishing, DNS poisoning experiments	ABB 800F, OpenPMC (PLC), Emerson MD, Turbogas Subsystem, Turbogas Control Subsystem, Steam cycle Subsystem Plant Control subsystem
6	Social Engineering	[16]	Testing	Simulation	Social engineering testing for enterprises using employee online data	Netkit
7	Cloud	[55]	Experiment	Emulation	DDoS attack testing on different network toplogies	OPENNEBULA, Netflow, Low Orbit Ion Canon
8	Autonomous System	[13]	Testing	Simulation	Military autonomous vehicle DDoS attack testing	JAUS messages, JSONS, NOSQL, PYTHON, RUBY, NODE.JS, JAVASCRIPT,XML, REST FULL WEBAPI

Table 2: Scenarios and there purpose in different domains

4.2.2 Monitoring

In this section we will talk about the methods, dashboard, layer and tools that are used for monitoring of cyber ranges and security test beds. Works in [4, 50, 98, 115, 19] use different data collection and analysis module for monitoring

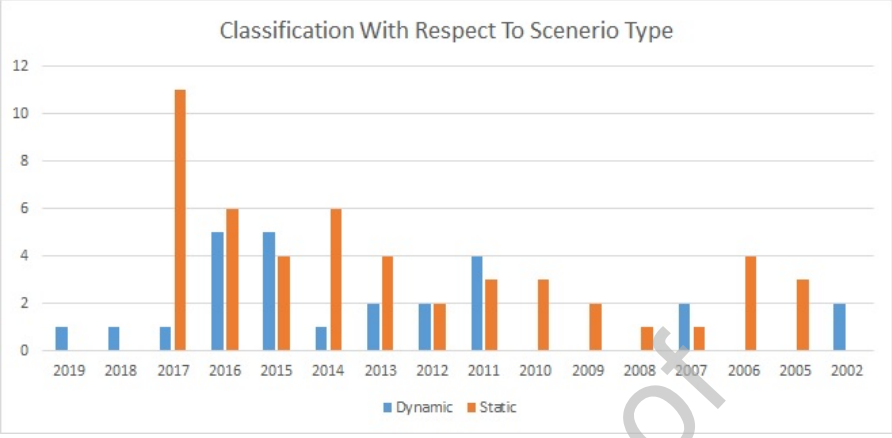

Figure 5: Classification of cyber-ranges and security testbeds based upon the scenario type which they support

purposes. While [108, 61, 65, 21, 39, 2, 71] use event logging mechanism and analysis techniques for monitoring purposes. [50, 115, 65, 21, 2] have specialized dashboards preset in the architecture to present the monitored information. [108, 61, 65, 21, 39, 2, 71] use mainly application layer protocols for data collection, while in [4, 50, 98, 115, 19] authors use network layer protocols for monitoring purposes. In term of tools these cyber ranges and security testbeds uses multitude of different tools, a detailed list of those tools is provided in section 4.4.5.

4.2.3 Learning

In this section we will discuss the learning and tutoring component, the after action analysis mechanism and scoring techniques present in different cyber ranges and security testbeds. Authors in [54, 34, 106, 117] have a tutoring or learning management system present in their functional architecture. These tutoring systems mainly consists of text, images and multimedia clips. Authors in [4] have an after action analysis module that operates over the complete experimental data set. Its main attribution is data pre-processing and calculation of a supplemental set of metrics derived from experimental bulk data. In term of scoring mechanisms, the work in [111] uses a score bot that is responsible for monitoring the status of the services and calculate the score for each team. While [34] use a scoreboard in which progress of participants is presented based upon the task the completed. Details of scoring mechanism and tools are presented in section 4.4.8.

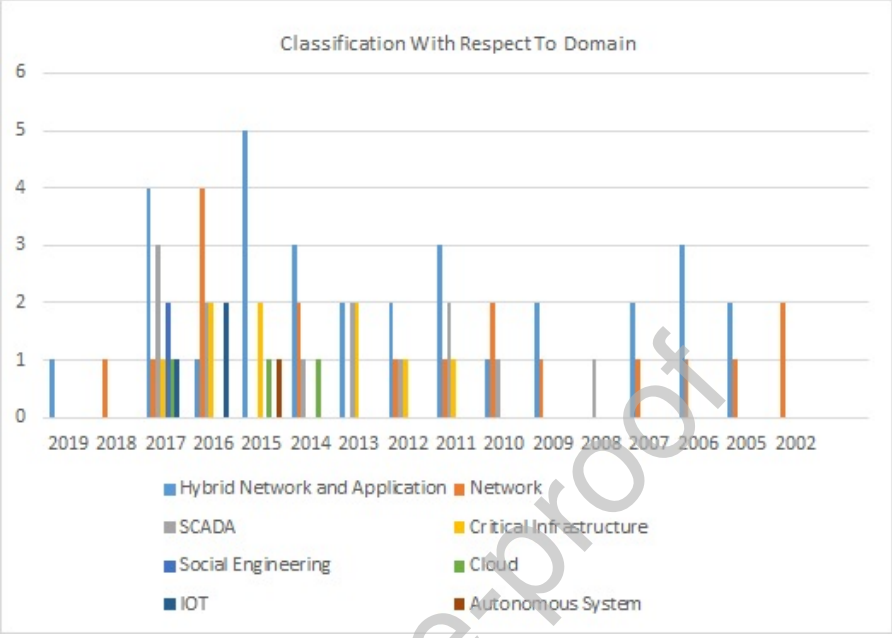

Figure 6: Classification of cyber-ranges and security testbeds based upon the scenarios domains

4.2.4 Management

In this section we present the roles, interfaces, range management, command and control, and resource management within the reviewed cyber ranges. Different teams perform different roles within the cyber range and security testbeds, we shared the details of different teams in section 4.2.5. In term of interfaces, [108, 5, 19] have dashboards that graphically presents the current state of cyber range and security test beds; while [2, 67] have special portals for communication. For interfaces, the work in [98] has a reporting module that is responsible for starting, enrolling devices and simulating. Authors in [117] have a remote desktop component that is used to initialize, start, monitor, and terminate remote desktop connections to machines. The work in [19] uses an API to manage remote access between different components of a cyber range, and authors in [44] use a proxy that enables running remote code and integrate different physical components. [71] have a control component that represents the main command and control for all the resources and services present with in the security test bed. The works in [90, 39, 111, 78, 13, 115] have data storage modules which stores elements like scenario models, attack tools, exercise and experiment rules and results; while authors in [2] have a module for cataloging different attack and defense scenarios.

4.2.5 Teaming

Figure 7 presents the different types of teams that participate in activities conducted at cyber ranges and security testbeds. The main types of teams are, red, blue, white, green, and autonomous teams. Red and blue teams, which corresponds to red and/or blue exercise types. Autonomous teams, in which there is the activities of a team is performed by an autonomous system, or agent, has gained an attention since 2014. Autonomous teams are added as a separate type to study the status of using automation of different team roles in cyber security exercises.

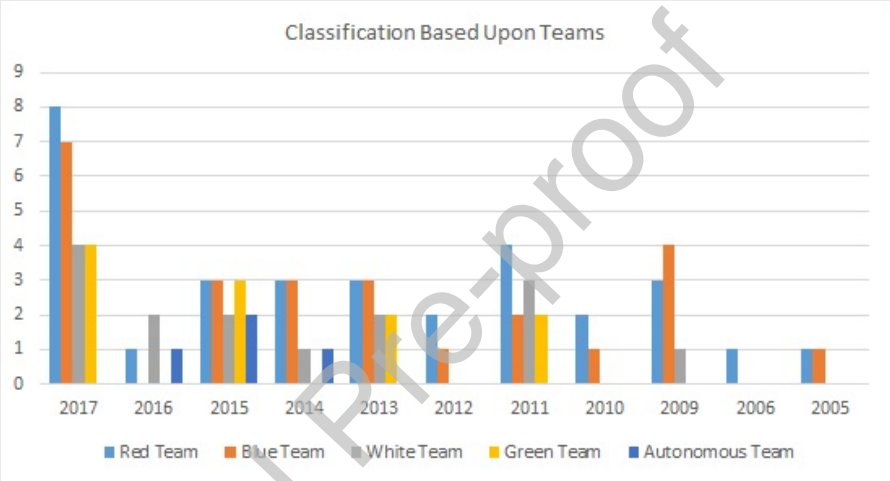

Figure 7: Classification of cyber-ranges and security testbeds based upon the teams

4.2.6 Environment

In this section we discuss the concept *environment*. This include the scenario execution environment type and different event generation tools that are used in scenario environments. Works in [34, 106, 37, 100, 23, 15, 117, 118, 4, 111] use an emulated environment for scenario execution. Their scenarios usually contain virtualized nodes running different services. Authors in [71, 13, 103, 81, 45, 21, 56, 61, 101, 75, 6, 19, 108] use Hybrid environment for the execution of cyber security scenarios. The environment contains both hardware, virtualized and simulated elements. The hardware usually contain specialized equipment, like PLCs that are difficult to emulate. In term of hardware based environments, works in [73, 65, 69, 7, 88, 90, 5, 39, 42] use actual hardware cyber security scenario execution, these scenarios are mostly relate to IoT, SCADA and critical infrastructure. Works in [44, 115, 10, 56, 66, 67, 106, 8] use different simulation and modeling techniques for cyber security scenario execution. Details of

different event generation tools, like traffic and user behavior, are presented in section 4.4.6 and 4.4.7.

Figure 8 indicates the type of the run time environment that are used in cyber ranges and security testbeds in the last 15 years. It can be sees that HW-only equipment has not been used widely. From 2002 until 2015, there has been only one paper presented a pure HW run time environment. Emulation has been, and still, used widely in cyber ranges and security testbeds. Since 2016, hybrid approaches have also become widely used.

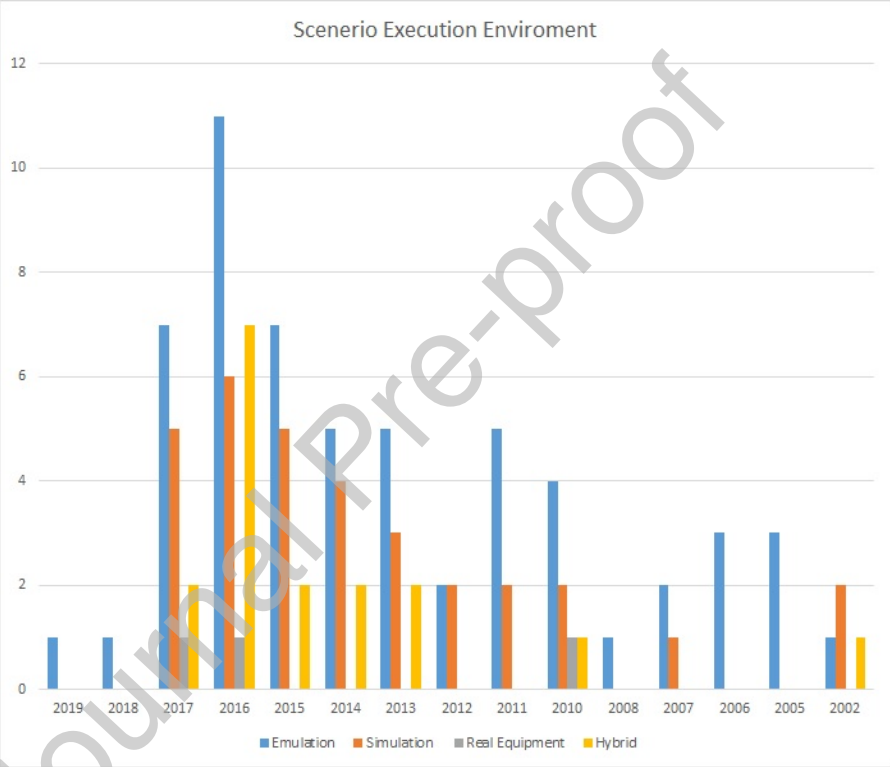

Figure 8: Classification of cyber-ranges and security testbeds based upon the scenario execution environment

4.3 Evaluation

In this section we discuss the different methods that have been used in order to evaluate cyber ranges and security testbeds. Out of 100 papers, 8 have details about the evaluation techniques employed in the cyber ranges and security testbed. Four papers used quantitative evaluation methods to evaluate the cyber ranges and security testbeds as a whole. The other four used qualitative

methods to evaluate the functionality of cyber ranges and security testbed by executing specific tests on them.

4.3.1 Overall and performance Evaluation

The following papers applied quantitative evaluation methods to evaluate the cyber ranged and security testbeds as a whole, especially the performance.

1. Researchers in [50] based their evaluation on the time for testbed generation. They measured the time required for generating an infrastructure of 3 router, 1 switch and 4 PCs' for an educational scenario. The total time required for generating the testbed was 42min 32s.

2. Researchers in [121] applied similar method and found out that the network environment generation tool took about 1624s to construct an environment consisting of three segments, i.e., the client, internal-server, and DMZ segments. For a single team in the cyber security exercise, there were five instances in total for each segment: the firewall, Windows 7 client, file server, database, and DNS/mail instances. It took about 6754s to finish the construction of identical segments for four teams for the conducted cyber security exercise.

3. In a distributed system scenario in [115], the researchers used Netbed's batch system to evaluate every possible combination of 7 bandwidths, 5 latencies, and 3 application parameter settings on four different configurations on a set of 20 nodes. The result was performing a total of 420 different tests in 30 hours, averaging 4.3 minutes each.

4. In simulation environment for validating protocols for distributed applications, researchers in [10] employed similar quantitative evaluation methods, which is also based upon time requirements.

4.3.2 Functional Evaluation

The following papers applied qualitative evaluation methods to evaluate the functionality of cyber ranges and security testbeds.

1. In a scenario of critical infrastructure protection [76], researchers employed CSET (cyber security evaluation tool) [1]. CSET is a qualitative evaluation method in which multiple security standards are integrated like NIST, Transportation Security Administration (TSA), North American Electric Reliability Corporation (NERC), U.S. Department of Defense (DoD), and others. When a security level is selected for evaluation, the CSET present a questionnaire based upon the above standards and measure the security level based upon the answers from security experts.

[1] https://ics-cert.us-cert.gov/Assessments

2. In another scenario of SCADA testbed and security device [56], researchers developed there own evaluation matrices for evaluating the security of SCADA testbed. Their evaluation matrices consist of.

    - The level of exposure of SCADA systems.
    - Ports of which the access is available (such as TCP/IP, MODBUS).
    - Access to websites connected with the SCADA system.
    - Vulnerabilities of websites connected with the SCADA system.
    - Vulnerabilities of Remote Terminal Unit(RTU) and Master Terminal Unit(MTU).
    - The status of common firewalls.

3. Researchers in a testbed of wearable IoT devices [98] employed a scenario based evaluation in which they determined what type of scenario capabilities their testbed supports. Scenario based evaluation takes into account the following capabilities in a scenario.

    - Scanning (e.g., IP and port scanning)
    - Fingerprinting
    - Process enumeration
    - Data leakage
    - Side-channel attacks
    - Data collection
    - Management access
    - Breaking encrypted traffic
    - Spoofing/masquerade attack
    - Communication delay attacks
    - Communication tampering
    - List known vulnerabilities
    - Vulnerability scan

4. In a cloud-based testbed for simulation of cyber attacks [58], researcher used two experiments to evaluate the testbed in a qualitative manner, in which they used slowHttptest to validate the effectiveness of a security module on a web server. In the first experiment a web server is equipped with a security module to mitigate a cyber attack, while in the second experiment a web server is targeted without the security module. During the first experiment the server became unavailable after 14 seconds of the attack. However, as soon as the duration of the connection reached the timeout set by the mitigation module, the connection was terminated and the server returned HTTP code 400. In the second experiment the server became unavailable after 14 seconds and remained in this state for next 586 seconds until the attack ended, as no mitigation module was activated.

4.4 Tools

In this section we identify and classify hardware and software tools utilized within contemporary cyber ranges and security test beds. Details of the tools with respect to year and domain of application as indicated in section 3.5 will be presented.

4.4.1 Emulation Tools

ID	Tool Name	Year	Paper	Domain
1	LAAS Cloud infrastructure	2014	[58]	Cloud
2	Openstack	2017	[31]	Cloud
3	EMULAB	2012	[44]	Critical Infrastructure
4	Unity Pro XL v7.0 suite	2015	[73]	Critical Infrastructure
5	EMULAB	2014	[95]	Critical Infrastructure
6	Virtual Box	2013	[105]	Critical Infrastructure
7	NetEm	2017	[120]	Critical Infrastructure
8	User-Mode Linux (UML)	2006	[54]	Hybrid Network and Application
9	Vmware Vsphere	2017	[17]	Hybrid Network and Application
10	Emulab	2015	[103]	Hybrid Network and Application
11	KVM	2016	[82]	Hybrid Network and Application
12	XEN Worlds	2010	[45]	Hybrid Network and Application
13	CITRIX XEN	2019	[21]	Hybrid Network and Application
14	Virtual Box	2015	[101]	Hybrid Network and Application
15	Vmware	2005	[51]	Hybrid Network and Application
16	Vmware	2011	[30]	Hybrid Network and Application
17	OPENNEBULA	2015	[19]	Hybrid Network and Application
18	OPENNEBULA	2015	[113]	Hybrid Network and Application
19	Qemu	2012	[117]	Hybrid Network and Application
20	KVM	2012	[117]	Hybrid Network and Application
21	XEN	2010	[23]	Hybrid Network and Application
22	OPEN VZ	2010	[23]	Hybrid Network and Application
23	Qemu	2011	[118]	Hybrid Network and Application
24	KVM	2011	[118]	Hybrid Network and Application
25	Mininet	2015	[8]	Hybrid Network and Application
26	Virtualbox	2014	[111]	Hybrid Network and Application
27	Virtual Machine	2010	[69]	Hybrid Network and Application
28	Cyber Smart	2009	[72]	Hybrid Network and Application
29	Vmware	2007	[15]	Hybrid Network and Application
30	Vmware ESXI	2013	[90]	Hybrid Network and Application
31	Vmware	2005	[86]	Hybrid Network and Application
32	Vmware ESXI	2013	[89]	Hybrid Network and Application
33	OpenFlow switches (OVS)	2016	[37]	IOT
34	Vmware Vsphere	2016	[37]	IOT
35	Qemu system	2016	[37]	IOT
36	XEN with the xapi toolstack	2017	[50]	Network
37	KVM	2016	[121]	Network

38	Vmware ESXI	2016	[121]	Network
39	OPENNEBULA	2014	[55]	Network
40	Xen-VM	2016	[20]	Network
41	Fluxbox desktop through Guacamole	2016	[20]	Network
42	Emulab	2006	[12]	Network
43	XEN	2014	[75]	Network
44	XORP Router	2009	[66]	Network
45	Open VZ	2009	[66]	Network
46	Future internet test bed FITS	2016	[4]	Network
47	Emulab	2018	[108]	Network
48	Emulab	2011	[97]	Network
49	Proxmox VE	2016	[81]	SCADA
50	Mininet	2017	[6]	SCADA
51	CORE emulator	2013	[3]	SCADA
52	Vmware Esxi	2012	[110]	SCADA
53	Vyatta software routers	2012	[110]	SCADA

Table 3: Emulation tools used in cyber ranges and security test beds

4.4.2 Simulation Tools

ID	Tool Name	Year	Paper	Domain
1	QualNet	2015	[13]	Autonmous Systems
2	Simulink	2015	[59]	Critical Infrastructure
3	Digsilent Powerfactory	2013	[49]	Critical Infrastructure
4	Real-time digital simulator	2013	[49]	Critical Infrastructure
5	Simulink	2014	[95]	Critical Infrastructure
6	SCADASim	2013	[105]	Critical Infrastructure
7	ModelNet	2002	[115]	Network
8	Network Simulator	2002	[115]	Network
9	Arena	2007	[60]	Network
10	Opnet	2016	[20]	Network
11	QualNet	2016	[20]	Network
12	ns2	2016	[20]	Network
13	ns3	2016	[20]	Network
14	PRIME (Parallel Real-time Immersive network Modeling Environment)	2009	[66]	Network
15	iSSFNet	2005	[67]	Network
16	Opnet	2011	[71]	SCADA
17	PowerWorld	2011	[71]	SCADA
18	Matlab	2014	[35]	SCADA
19	Simulink	2014	[35]	SCADA
20	Truetime	2014	[35]	SCADA
21	CIROS 6.0	2016	[81]	SCADA

22	Digital I/O, Analog I/O	2008	[56]	SCADA
23	MODBUS IO	2013	[3]	SCADA
24	Opnet	2012	[110]	SCADA
25	Matlab	2013	[42]	SCADA
26	Smulink	2013	[42]	SCADA
27	Simulink	2016	[5]	SCADA
28	Matlab	2016	[5]	SCADA
29	SimHydraulics	2016	[5]	SCADA
30	OpenPlc	2016	[5]	SCADA

Table 4: Simulation tools used in cyber ranges and security test beds

4.4.3 Hardware

ID	Tool Name	Year	Paper	Domain
1	Allen Bradley RSLogix 5000	2011	[76]	Critical Infrastructure
2	L35E PLCs.	2011	[76]	Critical Infrastructure
3	Factory Talk View 5.0 HMI screens	2011	[76]	Critical Infrastructure
4	Phasor measurement units	2011	[76]	Critical Infrastructure
5	Phasor data concentrator	2011	[76]	Critical Infrastructure
6	Synchrophasor vector processor	2011	[76]	Critical Infrastructure
7	protection relays controllers	2011	[76]	Critical Infrastructure
8	substation GPS clock	2011	[76]	Critical Infrastructure
9	Omicron relay test	2011	[76]	Critical Infrastructure
10	calibration device	2011	[76]	Critical Infrastructure
11	Real Time Digital Simulator (RTDS)	2011	[76]	Critical Infrastructure
12	amplifiers	2011	[76]	Critical Infrastructure
13	PMUs	2011	[76]	Critical Infrastructure
14	Cisco 5510	2011	[76]	Critical Infrastructure
15	MU Dynamics MU-4000 Analyzer	2011	[76]	Critical Infrastructure
16	IEEE C37.118,	2011	[76]	Critical Infrastructure
17	PLC	2015	[59]	Critical Infrastructure
18	Intelligebt End Device	2013	[49]	Critical Infrastructure
19	PLC	2013	[49]	Critical Infrastructure
20	PLC	2015	[41]	Critical Infrastructure
21	Remote Teminal Unit	2015	[41]	Critical Infrastructure
22	Smart Transmitter	2015	[41]	Critical Infrastructure
23	Cisco 6503	2014	[95]	Critical Infrastructure
24	IEC 60870-5- 104	2016	[47]	Critical Infrastructure
25	IEC 61850 MMS	2016	[47]	Critical Infrastructure
26	HP ProLiant DL380 G7	2015	[34]	Hybrid Network and Application
27	Google Glass	2016	[98]	IOT
28	Sony Smart watches	2016	[98]	IOT

29	Energy Managmement System	2018	[65]	IOT
30	Remote Terminal Unit	2018	[65]	IOT
31	Smart surveillance camera	2017	[39]	IOT
32	Android Smart Phone	2017	[39]	IOT
33	Cisco routers	2017	[50]	Network
34	Cisco routers	2010	[25]	Network
35	Siemens Devices	2010	[38]	SCADA
36	Emerson Devices	2010	[38]	SCADA
37	ABB Devices	2010	[38]	SCADA
38	Filed Dev	2010	[38]	SCADA
39	PLC	2017	[29]	SCADA
40	PLC	2016	[81]	SCADA
41	SIEMENS S7-300	2016	[81]	SCADA
42	Cisco ASA	2016	[81]	SCADA
43	RS485 Multiport	2008	[56]	SCADA
44	Phasor Data Concentrator	2016	[7]	SCADA
45	Phasor Measurement Units	2016	[7]	SCADA
46	SEL 421	2016	[7]	SCADA
47	Multifunction protection relays (7SJ610, 7SJ82)	2016	[7]	SCADA
48	SICAM PAS	2016	[7]	SCADA
49	Power TG	2016	[7]	SCADA
50	PLC	2013	[3]	SCADA
51	PLC	2016	[88]	SCADA
52	Raspbery PI	2016	[88]	SCADA
53	Cisco 2600 router	2012	[110]	SCADA
54	Juniper M61	2012	[110]	SCADA
55	PLC	2013	[42]	SCADA
56	Remote Teminal Unit	2013	[42]	SCADA
57	Rasbery PI	2016	[5]	SCADA

Table 5: Hardware devices used in cyber ranges and security test beds

4.4.4 Management Tools

1	Energy Management System	2013	[49]	Critical Infrastructure
2	Energy Management System	2014	[95]	Critical Infrastructure
3	Energy Management System	2016	[47]	Critical Infrastructure
4	ISEAGE	2013	[89]	Hybrid Network and Application
5	SIGAR API	2019	[21]	Hybrid Network and Application
6	3vilSh3llfor backdoor	2011	[30]	Hybrid Network and Application
7	vmService	2012	[117]	Hybrid Network and Application
8	vmService	2011	[118]	Hybrid Network and Application
9	HAMIDS	2017	[6]	SCADA
10	Xentop	2014	[75]	Network

Table 6: Management tools used in cyber ranges and security test beds

4.4.5 Monitoring Tools

1	Netflow	2014	[58]	Cloud
2	IPFIX	2014	[58]	Cloud
3	IPFIX	2017	[31]	Cloud
4	OSISoft PI Historian	2011	[76]	Critical Infrastructure
5	Zabbix	2012	[44]	Critical Infrastructure
6	Libpcap	2015	[59]	Critical Infrastructure
7	OSIsoft	2015	[59]	Critical Infrastructure
8	Wireshark	2015	[73]	Critical Infrastructure
9	Energy Management System	2013	[49]	Critical Infrastructure
10	Open V Switch	2015	[41]	Critical Infrastructure
11	Energy Management System	2014	[95]	Critical Infrastructure
12	Energy Management System	2016	[47]	Critical Infrastructure
13	Tcpdump	2017	[120]	Critical Infrastructure
14	Security Onion Linux	2017	[17]	Hybrid Network and Application
15	OSSEC	2017	[17]	Hybrid Network and Application
16	Tcpdump	2016	[82]	Hybrid Network and Application
17	Wireshark	2016	[82]	Hybrid Network and Application
18	SIGAR API	2019	[21]	Hybrid Network and Application
19	3vilSh3llfor backdoor	2011	[30]	Hybrid Network and Application
20	Nagios	2015	[19]	Hybrid Network and Application
21	Nagios	2015	[113]	Hybrid Network and Application
22	vmService	2012	[117]	Hybrid Network and Application
23	vmService	2011	[118]	Hybrid Network and Application
24	Catbird	2015	[8]	Hybrid Network and Application
25	ISEAGE	2013	[90]	Hybrid Network and Application
26	Snort	2005	[86]	Hybrid Network and Application
27	SyscallAnomaly	2005	[86]	Hybrid Network and Application
28	ISEAGE	2013	[89]	Hybrid Network and Application
29	Wireshark	2016	[98]	IOT
30	ADB	2016	[98]	IOT
31	Open V Switch	2016	[37]	IOT
32	Opendaylight controller	2016	[37]	IOT
33	Tcpdump	2017	[50]	Network
34	Tcpdump	2002	[115]	Network
35	Traceroute	2002	[115]	Network
36	FRONTIER	2010	[25]	Network
37	SHINE	2010	[25]	Network
38	Netflow	2014	[55]	Network
39	IPFIX	2014	[55]	Network
40	Emulab	2006	[12]	Network
41	Network Flight Recorder (NFR) Sentivist	2006	[12]	Network
42	FloodWatch	2006	[12]	Network
43	OPENFLOW	2014	[75]	Network
44	Xentop	2014	[75]	Network

45	Tcpdump	2009	[66]	Network
46	Testbed@TWISC Monitor	2018	[108]	Network
47	NAGIOS	2005	[2]	Network
48	Zabbix	2011	[97]	Network
49	NetDecoder	2017	[29]	SCADA
50	CanAnalyzer	2017	[29]	SCADA
51	Open V Switch	2016	[81]	SCADA
52	Pf sense	2016	[81]	SCADA
53	SNORT	2016	[81]	SCADA
54	OSSEC	2016	[81]	SCADA
55	HAMIDS	2017	[6]	SCADA
56	Wireshark	2012	[110]	SCADA
57	Tcpdump	2012	[110]	SCADA

Table 7: Monitoring tools used in cyber ranges and security test beds

4.4.6 Traffic Generation Tools

ID	Tool Name	Year	Paper	Domain
1	Low Orbit Ion Canon	2014	[58]	Cloud
2	Modbus	2011	[76]	Critical Infrastructure
3	Events (GOOSE)	2011	[76]	Critical Infrastructure
4	Generic Object Oriented Substation	2011	[76]	Critical Infrastructure
5	DNP3	2011	[76]	Critical Infrastructure
6	EtherNet/IP	2011	[76]	Critical Infrastructure
7	ISAGE	2013	[49]	Critical Infrastructure
8	Open flow	2015	[41]	Critical Infrastructure
9	Modbus	2014	[95]	Critical Infrastructure
10	DNP3	2014	[95]	Critical Infrastructure
11	Modbus	2016	[47]	Critical Infrastructure
12	ISEAGE	2013	[90]	Hybrid Network and Application
13	Traffic Collector/Replayer	2013	[89]	Hybrid Network and Application
14	Printer	2016	[98]	IOT
15	SSH	2016	[98]	IOT
16	SNMP	2016	[98]	IOT
17	MicroWorks	2018	[65]	IOT
18	SSH	2017	[50]	Network
19	SNMP	2017	[50]	Network
20	Policy Enabled Agent	2010	[25]	Network
21	Low Orbit Ion Canon	2014	[55]	Network
22	Emulab	2006	[12]	Network
23	hydra	2018	[108]	Network
24	tfn2k	2018	[108]	Network
25	Modbus Rsim	2011	[71]	SCADA
26	MODBUS	2008	[56]	SCADA
27	DNP3	2016	[88]	SCADA

28	Modbus	2016	[88]	SCADA
29	Virtual Control System Environment	2012	[110]	SCADA

Table 8: Traffic generation tools used in cyber ranges and security test beds

4.4.7 User Behavior Generation Tools

ID	Tool Name	Year	Paper	Domain
1	AMICI	2015	[103]	Hybrid Network and Application
2	ConsoleUser	2015	[101]	Hybrid Network and Application
3	AutoIT	2016	[20]	Network
4	Netkit	2017	[16]	Social Engineering

Table 9: Use behavior generation tools used in cyber ranges and security test beds

4.4.8 Scoring Tools and Mechanisms

ID	Tool Name	Year	Paper	Domain
1	Task Based	2013	[105]	Critical Infrastructure
2	Score Bot	2005	[51]	Hybrid Network and Application
3	Jeopardy board	2014	[99]	Hybrid Network and Application
4	ICTF score board, Flags	2011	[30]	Hybrid Network and Application
5	ICTF score board, Flags	2010	[23]	Hybrid Network and Application
6	Score Bot	2014	[111]	Hybrid Network and Application
7	Flags	2006	[100]	Hybrid Network and Application

Table 10: Scoring mechanisms and tools used in cyber ranges and security test beds

4.4.9 Scenario Definition

ID	Tool Name	Year	Paper	Domain
1	XML	2015	[13]	Autonmous Systems
2	JSON	2015	[13]	Autonmous Systems
3	XML	2012	[44]	Critical Infrastructure
4	YAML	2016	[82]	Hybrid Network and Application
5	XML	2013	[85]	Hybrid Network and Application
6	XML	2012	[117]	Hybrid Network and Application
7	XML	2011	[118]	Hybrid Network and Application
8	XML	2017	[50]	Network
9	XML	2010	[25]	Network
10	Integration Markup Language (IML)	2010	[25]	Network

11	Policy Editor Tools	2010	[25]	Network
12	Policy negotiation tool	2010	[25]	Network
13	XML	2007	[60]	Network
14	XML	2016	[20]	Network
15	XML	2002	[87]	Network
16	JSON	2016	[4]	Network
17	Offense and Defense Toolbox	2018	[108]	Network

Table 11: Scenario definition mechanisms in cyber ranges and security test beds

4.4.10 Security Testing Tools

ID	Tool Name	Year	Paper	Domain
1	Juas Messages	2015	[13]	Autonmous Systems
2	Low Orbit Ion Canon	2014	[58]	Cloud
3	Ettercap	2011	[76]	Critical Infrastructure
4	Ettercap	2015	[73]	Critical Infrastructure
5	GunPG1	2006	[54]	Hybrid Network and Application
6	John-the-Ripper	2006	[54]	Hybrid Network and Application
7	Bit torrent	2012	[10]	Hybrid Network and Application
8	Kali Linux	2017	[17]	Hybrid Network and Application
9	PathTest	2015	[103]	Hybrid Network and Application
10	Iperf	2015	[103]	Hybrid Network and Application
11	FTK Imager	2011	[46]	Hybrid Network and Application
12	Zora	2011	[46]	Hybrid Network and Application
13	netcat	2011	[46]	Hybrid Network and Application
14	cron	2011	[46]	Hybrid Network and Application
15	hex editor	2011	[46]	Hybrid Network and Application
16	offensivecomputing.net	2011	[46]	Hybrid Network and Application
17	Helix Forensics Live Linux CD	2011	[46]	Hybrid Network and Application
18	WinHex	2011	[46]	Hybrid Network and Application
19	md5sum	2011	[46]	Hybrid Network and Application
20	FTK Imager	2011	[46]	Hybrid Network and Application
21	vxheaven.org	2019	[21]	Hybrid Network and Application
22	SlowHTTPTest	2019	[21]	Hybrid Network and Application
23	LOIC	2019	[21]	Hybrid Network and Application
24	John the ripper	2006	[100]	Hybrid Network and Application
25	SVED	2015	[101]	Hybrid Network and Application
26	ENCASE Enterprise	2014	[99]	Hybrid Network and Application
27	WireShark	2014	[99]	Hybrid Network and Application
28	IDA Pro	2014	[99]	Hybrid Network and Application
29	Volatility	2014	[99]	Hybrid Network and Application
30	Hex Workshop	2014	[99]	Hybrid Network and Application
31	PDF Dissector	2014	[99]	Hybrid Network and Application

32	One-class support vector machine (OCSVM)	2018	[65]	IOT
33	Low Orbit Ion Canon	2014	[55]	Network
34	Crimeware toolkits	2016	[20]	Network
35	Metasploit	2016	[20]	Network
36	Nmap	2016	[20]	Network
37	Symantec ManHunt	2006	[12]	Network
38	Nmap	2011	[71]	SCADA
39	Nmap	2008	[56]	SCADA
40	Nessus	2008	[56]	SCADA
41	Wireshark	2008	[56]	SCADA
42	WinHTTrack	2008	[56]	SCADA
43	Netcraft	2008	[56]	SCADA
44	Kartoo	2008	[56]	SCADA

Table 12: Security Testing tools used in cyber ranges and security test beds

4.5 Future research trends and directions

In order to analyze the future research trends and direction, we looked closely to all papers since 2016 and we briefly present their future work in this section, and discuss and summarize them in section 5.2.

1. Design of cyber warfare testbed [21].
   Tow main future direction were proposed, the first is using *OS container*, as they are lightweight and support a wide range of OSs. The second direction is focusing on *simulating human behavior* using agent based simulation toolkit.

2. Testbed@ TWISC: A network security experiment platform [108].
   The authors of this work foresee threefold future development. The first is using virtualization and *SDN (SW Defined Networks)* due to its high programmability capability. The second is *federation*, which is required to support large scale exercises. Particularly they planned to use VPLS(Virtual Private LAN Service). Finally, they planned to work on what they call *Software Defined Security* that aims at tackling the additional attack vector on virtualization.

3. Achieving reproducible network environments with INSALATA [50].
   Few Future directions were proposed by the authors. They mainly focus on extending the current capability, e.g., (2) better *monitoring and event collection*, and (1) more *realistic* network environment reproducibility. Furthermore, *efficient deployment* is another goal for the future.

4. Capability Detection and Evaluation Matrics for Cyber Security lab Exercises [17].
   The authors planned to extend the experiment setting and invite different

students to take part for the sake of *cross validation*. *Stability* to support large scale exercises were also planned.

5. Control frameworks in network emulation testbeds: A survey [109].
   Three main directions can be identified in this paper, which are (1) supporting more *realistic* scenarios, and (2) *visualization and analytics*.

6. Cybersecurity training in control systems using real equipment [29].
   Further work of this work includes the *educational* evaluation of the laboratory.

7. Design and implementation of cybersecurity testbed for industrial IoT systems [65].
   The main future direction of this work is to use the testbed to *test and evaluate new security technologies* to various critical infrastructure systems, e.g., next generation intelligent power control system.

8. Developing a capability to classify technical skill levels within a Cyber Range [61].
   One idea that were discussed is the development of an *intent capability* whereby the intent of the user can be predicted.

9. Experiment as a service [32].
   The main future direction discussed in this paper is the development of *sharable and validated models (scenarios)* of *realistic* environments to support *federation*.

10. Extending Our Cyber-Range CYRAN with Social Engineering Capabilities [16].
    The social media profiles didn't use any real employee photo due to privacy concerns this can be improved in future using alternate images of employees. The content posted on social media is only text based in future other media formats like videos and images can be integrated for better representation of real social media.

11. Gamifying ICS security training and research: Design, implementation, and results of S3 [6].
    Future work discussed was to use the method applied in the paper as a foundation to enable others to run similar security *educational* experiments. This implies also the possibility to *share* the experiment models among different parties.

12. Improving and Measuring Learning Effectiveness at Cyber Defense Exercises [70].
    Future work was planned to develop a *learning* metrics and trends *benchmark*, which will provide a baseline to evaluate *learning* improvement in cybersecurity exercises.

13. KYPO Cyber Range: Design and Use Cases [113].
The future direction for KYPO is to use the current developed infrastructure to *test and experiment* with recent complex cyber attacks in order to evaluate and study *detection and mitigation* control against cyber threats to the critical infrastructure.

14. Modeling and simulation architecture for training in cyber defence education [106].
There are several courses of future development arising from the ideas presented above. A further direction is to make a comparison between our proposed architecture and existing military or commercial training solutions.

15. The FUSE testbed: establishing a microgrid for smart grid security experiments [120].
Similar to the previous future work, FUSE testbed was planned to be used to study methods and techniques to *detect* anomalies against critical infrastructure. *Security, availability and reliability* will be evaluated in the testbed to enhance *situational-awareness*.

16. Advanced security testbed framework for wearable IoT devices [98].
Afte completing the development of the testbed, the main future work discussed for this paper is to use the testbed in *testing* smart city IoT devices. The development of a lightweight anitmalware is also planned.

17. Alfons: A Mimetic Network Environment Construction System [121].
Optimizing and enhancing efficiency of the system are the main future work planned for the Alfons system.

18. Cybervan: A cyber security virtual assured network testbed [20].
In the following are the future work directions discussed for Cybervan: (1) *Scalability*, (2) *portability* to various virtualization and container technologies, (3) supporting more *realistic* scenarios (4) introducing *cognitive factors in simulation* of user/attacker behaviors, (5) enhancing *testing and validation* procedures of new technologies by developing an automated state space *exploration* mechanisms, and finally (6) enhancing *automation* capabilities in order to increase resource and research productivity.

19. CyRIS: A cyber range instantiation system for facilitating security training [82].
Two main issues were planned for future work of CyRIS system, the first is *scalability* and the second is *automation* of network configuration capabilities.

20. Design and architecture of an industrial IT security lab [81].
The two main directions planned for this work are to (1) apply the infrastructure for *education and awareness* training, and (2) perform advanced security *monitoring* by including remote production sites.

21. Developing a distributed software defined networking testbed for IoT [37].
The main future work discussed in this paper is to expand *simulation* capabilities to include IPv6 and evaluate performance evaluation.

22. PowerCyber: A remotely accessible testbed for Cyber Physical security of the Smart Grid [7].
The activities planned as future work include (1) developing *library of models* and datasets, (2) increasing the *user community*, and (3) developing advanced *realistic* use cases.

23. RIO: A denial of service experimentation platform in a Future Internet Testbed [4].
The main future work is to work on *efficiency* by studying the impact of each step on the experimentation overall time. Furthermore, the authors were planning to investigate possible *automation* of the platform.

24. Softgrid: A software-based smart grid testbed for evaluating substation cybersecurity solutions [47].
Future directions discussed for this work are multifold. (1) Supporting distributed setups and *emulation*, (2) *testing and evaluation* of different security solution and attack vectors, and (3) supporting other SCADA *protocols*, are the main directions discussed.

25. Virtualization of industrial control system testbeds for cybersecurity [5].
The future work presented focused on improving the studied *emulated* and virtual testbeds. Regarding virtualization, it was proposed to compare the system characteristics of both the virtual and the physical controller. Finally, *scalability* is the last issue the authors were planning to investigate.

5 Synthesis

The analysis of data related to tool yielded some interesting results. In term of scenario definition, XML is predominately used as indicated in table 11. XML provide a self descriptive way for designing and storing a scenario definition. The developed scenario definition can then be used in scenario simulation and emulation. It is used in autonomous systems, critical infrastructure, network and hybrid network and application scenarios. For monitoring, Tcpdump, IPFIX, and Wireshark were the most widely used tools in cyber ranges and security test-beds. They are used for monitoring traffic in cloud, network, critical infrastructure, and SCADA domains. Details of all the monitoring tools used in cyber ranges and security testbeds are presented in table 7. Multitude of different hardware devices were used in construction of different cyber ranges and security testbeds. However CISCO based devices are most widely used from the domain of critical infrastructure to networks and SCADA. Different PLC devices were also used in the construction of SCADA and critical infrastructure testbeds. Details of hardware devices used in construction of different cyber ranges and security testbeds are presented in table 5. For emulation,

Vmware based tools and Emulab were mostly used for critical infrastructure, hybrid network and application and networks domain. Vmware was also used in IoT and SCADA domains as well. Details of emulation tools used in cyber ranges and security testbeds is presented in table 3. In term of scenario simulation, Quanet, Simulink, Network Simulator and Matlab were widely used as indicated in table 4. Qualnet was used for both autonomous systems and critical infrastructure. Simulink was used for Critical infrastructure and SCADA. While Network Simulator and Matlab were used specifically for networks and SCADA respectively.

Different tools were used for traffic generation purposes in different domains. Modbus traffic is mostly used for SCADA and critical infrastructure while Low Orbit Ion Canon is used for TCP/UDP traffic generation. Details of traffic generation tools are presented in table 8. Different tools were used for security testing, user behavior generation, and scoring purposes in different domains, details of which are presented in tables 12, 9 10, respectively.

In term of the scenario types static and dynamic, a significant shift towards dynamic scenarios is witnessed in 2011 as indicated in figure 5. We believe that this shift happened due to identification of famous Stuxnet [62] worm in 2010 which created a lot of tidal waves in the cyber security research community. This observation is further backed by the data presented in figure 6, in which the rise of critical infrastructure and SCADA related testbeds can be observed. With the rise of cyber threats from nation state actors, investment in cyber security research increased with the aim to develop cyber resilience. This included development of new cyber security tools and methods as well as educating a workforce to handle cyber security crisis. This shift of sudden rise of education related scenarios in 2011 can be observed in figure 4. In the future, we believe that with abundant availability of computational resources, more and more tasks within the cyber ranges and security testbeds will get automated. From scenario creation to scenario execution and analysis, human role will become limited. This trend has started from 2014 with the appearance of autonomous teams in cyber ranges and security tetsbeds as indicated in figure 7.

5.1 Architecture and capabilities

In sections 4.2 and 4.1, we presented a new taxonomy that included general capabilities of cyber ranges and security testbeds. We also looked into the details of the architectural model of each cyber range reviewed in this paper. Analyzing the various architectural models of different systems we can see that the same components are named differently in different systems. For example, scenario execution element, orchestration module, controlling component could indicate the same functional component. We highlighted in section 4.2 the main concepts and used unified terminology. In this section, we aim at developing a unified functional architecture for cyber range based on the knowledge we gained from analyzing the architectures of cyber ranges and security testbeds, aforementioned.

Figure 9 shows a unified functional architecture that is developed from study-

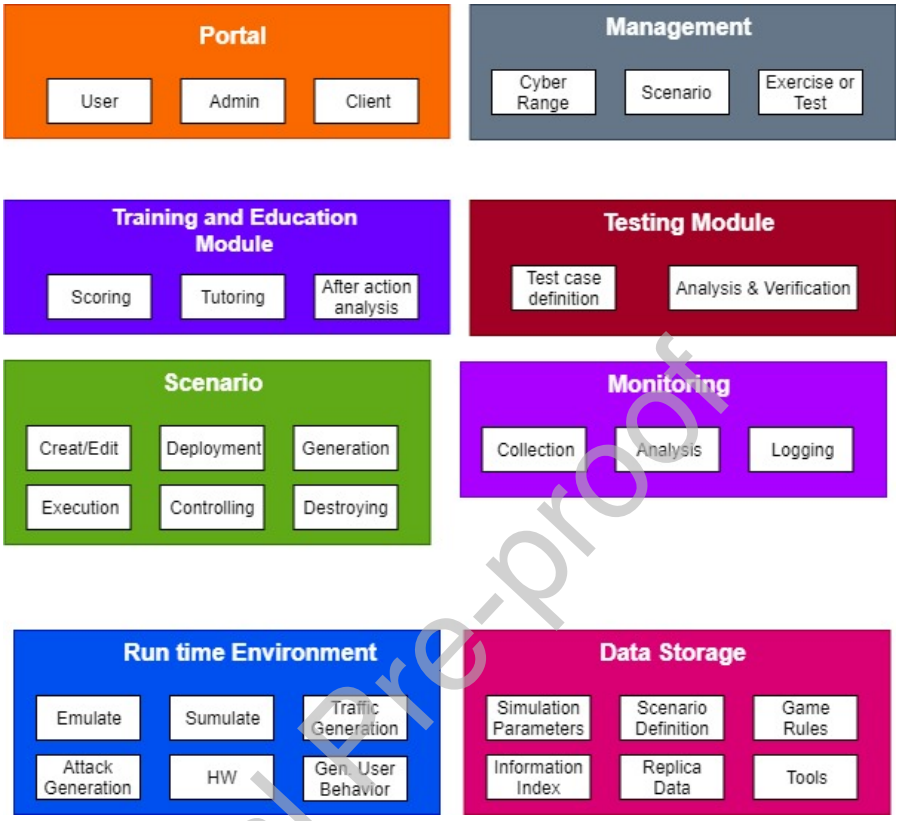

Figure 9: Cyber range and security testbed functional architecture

ing the literature. The architecture is divided into main components and within each component we define a set of sub-components.

- Portal
  Portal provide the interface for communication between the cyber range and security testbed to multiple users. The users can be cyber range admins, white team users to create and edit cyber security scenarios and other clients who use the cyber ranges for various tests and experiments. The cyber range and security testbed admin user performs over management activities related to the cyber range or the testbed, which includes resource management and access management to other users like instructor, testers, trainee, or a white team member scenario creator. The scenario creator creates scenarios which can be deployed for cyber security exercise and experiments. The clients can use cyber range and testbed resources for testing and experimentation according to their requirements.

- Management

In management functions, resources and roles are managed. Resources includes the memory, processing and storage capabilities. While roles management include the assignment of duties for the cyber security exercises and experiment. Cyber range and security testbed management is related to overall range management, it deals with assigning roles to exercise and experiment managers, as well as necessary computational resources to conduct the exercise and run the experiment. Multiple exercises and experiments can be conducted at a same time on cyber range and security testbed. Exercise management deals with the segregation of roles and resource of an exercise or an experiment participant. In an exercise or experiment, multiple scenarios can be conducted, scenario management deals with the management of multiple exercises or experiment scenarios on the environment. Extensive collection of log information and analysis is performed from the cyber range and security testbed infrastructure for managing the cyber range and security testbed infrastructure in optimal manner.

- Training and education
  Training and education module provides tutoring system for cyber range and security testbed. The tutoring system consist of cyber security concepts and their practical exercises for cyber security education purposes. The training outcome is evaluated using a scoring mechanism. Multiple scoring mechanism can be used like flag-based scoring, task-based scoring and scoring with the help of event log information. After action analysis using training participant feedback and event information is performed to remove inefficiencies in conducting cyber security exercises and improve their qualities.

- Testing
  As mentioned before, besides training, the second main objective of a cyber range is testing and security assessment. We noticed two main types of tests that can be conducted in a cyber range. The first is to test the security of a system or a product, and the second is to test a new defence or attack method or technique. Testing module aims at defining the test cases, which will be turned into a scenario that will be deployed and executed on the run time environment. After executing the scenario, through the scenario module and the run time environment, the result will be sent back to the testing module to conduct the final analysis and evaluation of the system under test.

- Scenario
  White team members have access to scenario creator interface. The scenario creator interface is used to create, edit, deploy, generate, execute, control and destroy cyber security scenarios. The scenario creator gives capability to design and deploy new cyber security scenarios and save the scenario configuration in a file. The scenario editor allows to edit predefined scenarios for modification. The scenario deployer read the saved

scenario configuration file and deploy the scenario on emulated, simulated, or hybrid environment. The scenario generator is used to generate new cyber security scenario using minimum scenario configurations. The scenario executor executes the scenario and perform different actions during different phases of the scenario, like injecting network traffic or initiating a user behavior at different stages to make the scenario more realistic. The scenario controller gives the functionality of modifying the scenario during execution. The scenario destroyer used to remove obsolete scenario from cyber security exercises to be ready for the next exercise.

- Monitoring
  Monitoring provides the capability to monitor cyber security exercise and experiment execution. It includes collection of logs from multiple sources and analysis on those logs. The log sources contain different network and operating system interfaces. The logs are mostly in different formats, so their format needs to be unified using some pre-processing techniques. Analysis is then performed on the unified logs to identify different activities being performed by cyber security exercise and experiment participants at different stages of an exercise and an experiment scenario.

- Run time environment
  The run time environment represents the infrastructure layer that contains physical, virtual, hybrid and cloud platforms, on which the scenario is deployed. Red team attacks the infrastructure and blue team defends the infrastructure. The activities of both teams create events, which are used for monitoring and scoring purposes. To make the cyber security exercise and experiment environment more realistic, user behavior and random network traffic is generated.

- Data Storage
  Data storage aims at storing various artifacts needed for executing the training, or testing, scenarios. It includes scenario definition files, information about the rules that need to be implemented in the scenario, and tools required for the scenario execution. The data storage act as a library for the scenarios with relevant meta data related to scenario difficulty and complexity. This assists in designing cyber security exercise and experiment according to the skill set of participants.

- Teaming
  Although not presented in functional architecture of cyber range and security testbeds, teaming roles can't be ignored. A white team is responsible for scenario creation and setting the learning objectives for the scenario. Green team in involved in the monitoring of the scenario. While red and blue teams have access to run time environment for scenario execution. Autonomous teams can be used to emulate or simulate any role of red, blue, white and green teams.

5.1.1 Ideal methods and tools

In this section we will discuss about the ideal methods and tools for cyber range and security testbed development. First, we would like to argue that there is a lack in standards for cyber ranges and security testbed development. Their is a need to standardize this field, we found cyber range interoperability standards [27] that governs the federation principles for cyber ranges and security testbeds. So we would like to suggest any future development of cyber ranges and security should be governed by accepted standards. Secondly, from the results indicated in figure 6 it can be augured that hybrid network and application domain over emulation is most popular for cyber range and security testbed development. Therefore, we expect to see more research in the field. We would like to suggest the use of open source or publicly available tools for their development. For hybrid network and application domains, we would like to suggest use of cloud infrastructure like Opennebula or Openstack for emulation due to their standardize work environment. With cloud, we would also suggest the use of standard APIs for communication with specific hardware which can't be emulated like PLCs. APIs should also be used for management, monitoring and giving access to teams on the cyber range and security testbed.

5.2 Future research trends and directions

In section 4.5 we presented the main future plans for all recent work related to cyber range and security testbeds. In this section we compile these plans and provide the main directions for future work. We categorize the future direction into the following categories

1. Efficiency
   One of the main topics for future work that were discussed by reviewed papers is enhancing the efficiency of exercise lifecycle. To do that, automation is mentioned as a possible technique to make the deployment and execution of exercises more efficient [82, 14, 50].

2. Scalability, realism and virtualization
   To achieve the best result from a training exercise or a testing process, the run time environment should be as close as possible to the real world. While developing small scale and class-room oriented testbeds is feasible and easy to achieve, scaling the testbed to provide as realistic scenarios as possible is a challenging task. Scalability is mentioned by many papers as one of enhancement plans for their cyber ranges [20, 82, 14]. Using the new virtualization and emulation techniques, e.g., SDN, is put as an option. Particularly, SDN provides a high degree of programmability that is desired in such settings. Container technology and its support lightweight nature was another scalability enabled future technology. Regarding the issue of realism, one paper proposed to provide the support for a larger number of protocols, e.g., SCADA protocols, in future design of security testbeds [5].

3. Federation
   Another related topic is federation. Federation is also mentioned by couple of papers as one of the main future direction. Activities and issues related to federation include "sharability", potability, support of multiple locations, developing standard way to describe scenarios, defining a library for models and data, and expanding the user community [108, 32].

4. User behavior simulation
   Current work identified that techniques used today for user behavior simulation has its limitation. To overcome its limitation, advances in user behavior simulation is proposed as one potential future work [16, 21]. Examples of the proposed enhancements in the future are to use agent based simulations and introducing cognitive factors.

5. Monitoring
   Monitoring capabilities are essential for any cyber range or security testbed installations. However, the degree of monitoring and the way it can be used vary from one solution to another. Future work related to monitoring is to use advanced security monitoring and data collection techniques [50, 81].

6. Testing and evaluation
   Few papers proposed, as future work, to extend the current cyber ranges and security testbeds with new testing and evaluation capabilities in order to (1) test new security solutions and technologies [65], (2) testing new attack vectors and attack techniques [113], (3) testing for some security features that were not considered before in the testbeds like reliability and availability [47], and (4) enhancing the testing techniques [4].

7. Education and learning
   One of the issues that are missing in many current cyber ranges and training testbeds is considering learning and educational aspects [81]. Thus, future work was proposed to support techniques and methods to evaluate learning effectiveness and improvements, e.g., by developing learning metrics [17].

8. Benchmarking
   The final aspect that we identified as future work is the plans to conduct comparisons between the developed cyber ranges and security testbeds and others. In order to support this activities, we believe that developing a cyber range benchmark is essential for the future [106].

6 Discussion and Conclusion

From the systematic literature review we confirmed our observations that the interest in cyber range and security testbeds has increased in the last few years as indicated in figure 2. We identified that scenarios play a major role in cyber

range and security testbed development as indicated in figure 2. These scenarios focuses on cyber security testing, experimental and educational purposes as indicated in figure 4. These scenarios are executed on emulated, simulated, hybrid and real equipment environment as indicted in figure 8. The execution of theses scenarios is either static or dynamic as indicated in figure 5. Static scenarios have a linear execution and they execute according to predefined process. Dynamic scenarios have a non linear execution and their execution depends upon the dynamic changes that are introduced in to the environment. These Dynamic changes are introduced by the teams involved in the scenario.

Most of the uses cases of cyber ranges and security testbeds are centered around the needs for red and blue team training as indicated in figure 7. The role of white and green teams need to be focused for cyber security scenario development and cyber security scenario management. A new trend of autonomous teams is starting to appear in cyber ranges and security testbed. These teams automate the role of red, blue and white teams, to reduce the time required in conducting cyber security exercises, tests and experiments. However concrete methods to model the behavior of these teams are missing and *modeling of attack and defense scenarios for cyber ranges and security testbeds* are required for systemic execution and evaluation of a cyber security scenario.

The interest to use cyber ranges in testing, besides education, has increased in the last few years. This indicates that cyber ranges are not exclusively educational platforms, but can be used in other purposes, like testing. Most of the security test beds and cyber ranges are focusing on either quantitative evaluation methods or qualitative evaluation methods. Evaluation criteria which focuses on both quantitative and qualitative analysis on the security test bed and cyber ranges is missing. New evaluation metrics which focus on evaluating a single scenario on multiple test beds on both qualitative and quantitative manner will assist the evaluation of the security testbeds and cyber ranges in a systemic comparative analysis.

The figure 6 indicates that networking systems were the main application domain for cyber ranges and security testbeds, SCADA system started to gain attention from 2010, and in recent year cyber ranges and security testbeds have covered most application domain aforementioned. IOT, social engineering and testbeds for autonomous system are being developed. However most of these testbeds uses a hybrid environment in which the combine emulation, simulation and real equipment to produce most realistic cyber security environment for cyber security exercises, training, education and experiments.

References

[1] *A Comprehensive List of Cyber Security Competitions.* URL: `https://cybersecuritydegrees.com/faq/comprehensive-list-of-cyber-security-competitions/`.

[2] Roberto Alfieri et al. "The INFN-grid testbed". In: *Future Generation Computer Systems* 21.2 (2005), pp. 249–258.

[3] Abdulmohsen Almalawi et al. "SCADAVT-A framework for SCADA security testbed based on virtualization technology". In: *Local Computer Networks (LCN), 2013 IEEE 38th Conference on*. IEEE. 2013, pp. 639–646.

[4] Igor Drummond Alvarenga and Otto Carlos MB Duarte. "RIO: A denial of service experimentation platform in a Future Internet Testbed". In: *Network of the Future (NOF), 2016 7th International Conference on the*. IEEE. 2016, pp. 1–5.

[5] Thiago Alves, Rishabh Das, and Thomas Morris. "Virtualization of industrial control system testbeds for cybersecurity". In: *Proceedings of the 2nd Annual Industrial Control System Security Workshop*. ACM. 2016, pp. 10–14.

[6] Daniele Antonioli et al. "Gamifying ICS security training and research: Design, implementation, and results of S3". In: *Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and PrivaCy*. ACM. 2017, pp. 93–102.

[7] Aditya Ashok, Sujatha Krishnaswamy, and Manimaran Govindarasu. "PowerCyber: A remotely accessible testbed for Cyber Physical security of the Smart Grid". In: *Innovative Smart Grid Technologies Conference (ISGT), 2016 IEEE Power & Energy Society*. IEEE. 2016, pp. 1–5.

[8] Mahmoud Al-Ayyoub et al. "Sdsecurity: A software defined security experimental framework". In: *Communication Workshop (ICCW), 2015 IEEE International Conference on*. IEEE. 2015, pp. 1871–1876.

[9] D Balenson, L Tinnel, and T Benzel. "Cybersecurity experimentation of the future (CEF): catalyzing a new generation of experimental cybersecurity research". In: *SRI International, Tech. Rep.* (2015).

[10] Marinho P Barcellos et al. "Beyond network simulators: Fostering novel distributed applications and protocols through extendible design". In: *Journal of Network and Computer Applications* 35.1 (2012), pp. 328–339.

[11] Terry Benzel et al. "Current developments in DETER cybersecurity testbed technology". In: *Conference for Homeland security, 2009. CATCH'09. Cybersecurity applications & technology*. IEEE. 2009, pp. 57–70.

[12] Terry Benzel et al. "Experience with deter: a testbed for security research". In: *Testbeds and Research Infrastructures for the Development of Networks and Communities, 2006. TRIDENTCOM 2006. 2nd International Conference on*. IEEE. 2006, 10–pp.

[13] Dennis Lee Bergin. "Cyber-attack and defense simulation framework". In: *The Journal of Defense Modeling and Simulation* 12.4 (2015), pp. 383–392.

[14] Razvan Beuran et al. "Integrated Framework for Hands-on Cybersecurity Training: CyTrONE". In: *Computers & Security* (2018).

[15] Charles Border. "The development and deployment of a multi-user, remote access virtualization system for networking, security, and system administration classes". In: *ACM SIGCSE Bulletin* 39.1 (2007), pp. 576–580.

[16] Sam Braidley. "Extending Our Cyber-Range CYRAN with Social Engineering Capabilities". In: *De Montfort University, MSc Thesis Report* (Sept. 2016).

[17] Emin Caliskan et al. "Capability Detection and Evaluation Metrics for Cyber Security lab Exercises". In: *ICMLG2017 5th International Conference on Management Leadership and Governance*. Academic Conferences and publishing limited. 2017, p. 407.

[18] *Carnegie Mellon University - Software Engineering Institute*. URL: `https://stepfwd.cert.org/lms`.

[19] Pavel Čeleda et al. "KYPO–A Platform for Cyber Defence Exercises". In: *M&S Support to Operational Tasks Including War Gaming, Logistics, Cyber Defence. NATO Science and Technology Organization* (2015).

[20] Ritu Chadha et al. "Cybervan: A cyber security virtual assured network testbed". In: *Military Communications Conference, MILCOM 2016-2016 IEEE*. IEEE. 2016, pp. 1125–1130.

[21] Yogesh Chandra and Pallaw Kumar Mishra. "Design of Cyber Warfare Testbed". In: *Software Engineering*. Springer, 2019, pp. 249–256.

[22] C Jason Chiang et al. "Cyber Testing Tools and Methodologies". In: *Presentation at ITEA, November* (2013).

[23] Nicholas Childers et al. "Organizing large scale hacking competitions". In: *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer. 2010, pp. 132–152.

[24] Okoli Chitu and Schabram Kira. "A Guide to Conducting a Systematic Literature Review of Information Systems Research". In: *Sprouts: Working Papers on Information Systems* 26.10 (2010).

[25] Edward Chow et al. "An Intelligent network for federated testing of NetCentric systems". In: *Policies for Distributed Systems and Networks (POLICY), 2010 IEEE International Symposium on*. IEEE. 2010, pp. 44–52.

[26] Mehmet Hazar Cintuglu et al. "A Survey on Smart Grid Cyber-Physical System Testbeds." In: *IEEE Communications Surveys and Tutorials* 19.1 (2017), pp. 446–464.

[27] Suresh K Damodaran and Kathy Smith. *CRIS Cyber Range Lexicon, Version 1.0*. Tech. rep. Massachusetts Institute of Technology Lexington Lincoln Laboratory, 2015.

[28] Jon Davis and Shane Magrath. *A survey of cyber ranges and testbeds*. Tech. rep. Defence Science, Technology Organization Edinburgh (Australia) Cyber, and Electronic Warfare Division, 2013.

[29] Manuel Domınguez et al. “Cybersecurity training in control systems using real equipment”. In: *IFAC-PapersOnLine* 50.1 (2017), pp. 12179–12184.

[30] Adam Doupé et al. “Hit’em where it hurts: a live security exercise on cyber situational awareness”. In: *Proceedings of the 27th Annual Computer Security Applications Conference*. ACM. 2011, pp. 51–61.

[31] Thomas W Edgar and David O Manz. *Research Methods for Cyber Security*. Syngress, 2017.

[32] Thomas W Edgar and Theora R Rice. “Experiment as a service”. In: *Technologies for Homeland Security (HST), 2017 IEEE International Symposium on*. IEEE. 2017, pp. 1–6.

[33] Thomas Edgar, David Manz, and Thomas Carroll. “Towards an experimental testbed facility for cyber-physical security research”. In: *Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research*. ACM. 2011, p. 53.

[34] Margus Ernits, Johannes Tammekänd, and Olaf Maennel. “i-tee: A fully automated Cyber Defense Competition for Students”. In: *ACM SIGCOMM Computer Communication Review*. Vol. 45. 4. ACM. 2015, pp. 113–114.

[35] Adnan A Farooqui et al. “Cyber security backdrop: A scada testbed”. In: *Computing, Communications and IT Applications Conference (ComComAp), 2014 IEEE*. IEEE. 2014, pp. 98–103.

[36] Bernard Ferguson, Anne Tall, and Denise Olsen. “National cyber range overview”. In: *Military Communications Conference (MILCOM), 2014 IEEE*. IEEE. 2014, pp. 123–128.

[37] Olivier Flauzac, Carlos Gonzalez, and Florent Nolot. “Developing a distributed software defined networking testbed for IoT”. In: *Procedia Computer Science* 83 (2016), pp. 680–684.

[38] Igor Nai Fovino et al. “An experimental platform for assessing SCADA vulnerabilities and countermeasures in power plants”. In: *Human System Interactions (HSI), 2010 3rd Conference on*. IEEE. 2010, pp. 679–686.

[39] Angelo Furfaro et al. “Using virtual environments for the assessment of cybersecurity issues in IoT scenarios”. In: *Simulation Modelling Practice and Theory* 73 (2017), pp. 43–54.

[40] Steven Furnell, Pete Fischer, and Amanda Finch. “Can’t get the staff? The growing need for cyber-security skills”. In: *Computer Fraud & Security* 2017.2 (2017), pp. 5–10.

[41] Haihui Gao et al. “Cyber-physical systems testbed based on cloud computing and software defined network”. In: *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2015 International Conference on*. IEEE. 2015, pp. 337–340.

[42] Haihui Gao et al. “The design of ics testbed based on emulation, physical, and simulation (eps-ics testbed)”. In: *Intelligent Information Hiding and Multimedia Signal Processing, 2013 Ninth International Conference on*. IEEE. 2013, pp. 420–423.

[43] Anastasius Gavras et al. “Future internet research and experimentation: the FIRE initiative”. In: *ACM SIGCOMM Computer Communication Review* 37.3 (2007), pp. 89–92.

[44] Béla Genge, Christos Siaterlis, and Marc Hohenadel. “Amici: An assessment platform for multi-domain security experimentation on critical infrastructures”. In: *International Workshop on Critical Information Infrastructures Security*. Springer. 2012, pp. 228–239.

[45] Nathaniel Gephart and Benjamin A Kuperman. “Design of a virtual computer lab environment for hands-on information security exercises”. In: *Journal of Computing Sciences in Colleges* 26.1 (2010), pp. 32–39.

[46] Sonja M Glumich and Brian A Kropa. *DefEX: Hands-On Cyber Defense Exercise for Undergraduate Students*. Tech. rep. Air Force Research Lab Rome NY Information Directorate, 2011.

[47] Prageeth Gunathilaka, Daisuke Mashima, and Binbin Chen. “Softgrid: A software-based smart grid testbed for evaluating substation cybersecurity solutions”. In: *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. ACM. 2016, pp. 113–124.

[48] Rahul Gurnani, Kaushik Pandey, and Shashi Kant Rai. “A scalable model for implementing Cyber Security Exercises”. In: *Computing for Sustainable Global Development (INDIACom), 2014 International Conference on*. IEEE. 2014, pp. 680–684.

[49] Adam Hahn et al. “Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid”. In: *IEEE Transactions on Smart Grid* 4.2 (2013), pp. 847–855.

[50] Nadine Herold et al. “Achieving reproducible network environments with INSALATA”. In: *IFIP International Conference on Autonomous Infrastructure, Management and Security*. Springer. 2017, pp. 30–44.

[51] Lance J Hoffman et al. “Exploring a national cybersecurity exercise for universities”. In: *IEEE Security & Privacy* 3.5 (2005), pp. 27–33.

[52] Hannes Holm and Teodor Sommestad. “Sved: Scanning, vulnerabilities, exploits and detection”. In: *Military Communications Conference, MILCOM 2016-2016 IEEE*. IEEE. 2016, pp. 976–981.

[53] Hannes Holm et al. “A Survey of Industrial Control System Testbeds”. In: *Secure IT Systems*. Ed. by Sonja Buchegger and Mads Dam. Cham: Springer International Publishing, 2015, pp. 11–26. ISBN: 978-3-319-26502-5.

[54] Ji Hu, Dirk Cordel, and Christoph Meinel. “A virtual machine architecture for creating IT-security laboratories”. In: (2006).

[55] Tomas Jirsik et al. "Cloud-based security research testbed: A DDoS use case". In: *Network Operations and Management Symposium (NOMS), 2014 IEEE*. IEEE. 2014, pp. 1–2.

[56] Sungmo Jung, Jae-gu Song, and Seoksoo Kim. "Design on SCADA testbed and security device". In: *International Journal of Multimedia and Ubiquitous Engineering* 3.4 (2008), pp. 75–86.

[57] Jason Kick. *Cyber exercise playbook*. Tech. rep. MITRE CORP BEDFORD MA, 2014.

[58] Daniel Kouril et al. "Cloud-based testbed for simulation of cyber attacks". In: *Network Operations and Management Symposium (NOMS), 2014 IEEE*. IEEE. 2014, pp. 1–6.

[59] Georgia Koutsandria et al. "A real-time testbed environment for cyber-physical security on the power grid". In: *Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy*. ACM. 2015, pp. 67–78.

[60] Michael E Kuhl et al. "Cyber attack modeling and simulation for network security analysis". In: *Proceedings of the 39th Conference on Winter Simulation: 40 years! The best is yet to come*. IEEE Press. 2007, pp. 1180–1188.

[61] William Aubrey Labuschagne and Marthie Grobler. "Developing a capability to classify technical skill levels within a Cyber Range". In: *ECCWS 2017 16th European Conference on Cyber Warfare and Security*. Academic Conferences and publishing limited. 2017, p. 224.

[62] Ralph Langner. "Stuxnet: Dissecting a cyberwarfare weapon". In: *IEEE Security & Privacy* 9.3 (2011), pp. 49–51.

[63] Sylvain P. Leblanc et al. "An Overview of Cyber Attack and Computer Network Operations Simulation". In: *Proceedings of the 2011 Military Modeling & Simulation Symposium*. MMS '11. Boston, Massachusetts: Society for Computer Simulation International, 2011, pp. 92–100. URL: http://dl.acm.org/citation.cfm?id=2048558.2048572.

[64] Sylvain P Leblanc et al. "An overview of cyber attack and computer network operations simulation". In: *Proceedings of the 2011 Military Modeling & Simulation Symposium*. Society for Computer Simulation International. 2011, pp. 92–100.

[65] Seokcheol Lee et al. "Design and implementation of cybersecurity testbed for industrial IoT systems". In: *The Journal of Supercomputing* (2017), pp. 1–15.

[66] Yue Li, Michael Liljenstam, and Jason Liu. "Real-time security exercises on a realistic interdomain routing experiment platform". In: *Principles of Advanced and Distributed Simulation, 2009. PADS'09. ACM/IEEE/SCS 23rd Workshop on*. IEEE. 2009, pp. 54–63.

[67] Michael Liljenstam et al. “Rinse: The real-time immersive network simulation environment for network security exercises”. In: *Proceedings of the 19th Workshop on Principles of Advanced and Distributed Simulation*. IEEE Computer Society. 2005, pp. 119–128.

[68] Maria B Line and Nils Brede Moe. “Understanding collaborative challenges in it security preparedness exercises”. In: *IFIP International Information Security Conference*. Springer. 2015, pp. 311–324.

[69] George Louthan et al. “The Blunderdome: An Offensive Exercise for Building Network, Systems, and Web Security Awareness.” In: *CSET*. 2010.

[70] Kaie Maennel, Rain Ottis, and Olaf Maennel. “Improving and Measuring Learning Effectiveness at Cyber Defense Exercises”. In: *Nordic Conference on Secure IT Systems*. Springer. 2017, pp. 123–138.

[71] Malaz Mallouhi et al. “A testbed for analyzing security of SCADA control systems (TASSCS)”. In: *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*. IEEE. 2011, pp. 1–7.

[72] Jim Marshall. “The Cyber Scenario Modeling and Reporting Tool (CyberSMART)”. In: *Cybersecurity Applications & Technology Conference For Homeland Security*. IEEE. 2009, pp. 305–309.

[73] Estefanıa Etchevés Miciolino et al. “Communications network analysis in a SCADA system testbed under cyber-attacks”. In: *Telecommunications Forum Telfor (TELFOR), 2015 23rd*. IEEE. 2015, pp. 341–344.

[74] Jelena Mirkovic et al. “The DETER project: Advancing the science of cyber security experimentation and test”. In: *Technologies for Homeland Security (HST), 2010 IEEE International Conference on*. IEEE. 2010, pp. 1–7.

[75] Igor M Moraes et al. “FITS: A flexible virtual network testbed architecture”. In: *Computer Networks* 63 (2014), pp. 221–237.

[76] Thomas Morris et al. “A control system testbed to validate critical infrastructure protection concepts”. In: *International Journal of Critical Infrastructure Protection* 4.2 (2011), pp. 88–103.

[77] Julien Murphy et al. “Building a Virtual Cybersecurity Collaborative Learning Laboratory (VCCLL)”. In: *Proceedings of the International Conference on Security and Management (SAM)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). 2014, p. 1.

[78] Radek Ošlejšek et al. “Towards a Unified Data Storage and Generic Visualizations in Cyber Ranges”. In: *ECCWS 2017 16th European Conference on Cyber Warfare and Security*. Academic Conferences and publishing limited. 2017, p. 298.

[79] Amanda Palleschi. “Pentagon fought proposal: Congress Adopts Provision To Halt Funding For National Cyber Range”. In: *Inside the Air Force* 21.51 (2010), pp. 10–10.

[80] Victor-Valeriu Patriciu and Adrian Constantin Furtuna. "Guide for designing cyber security exercises". In: *Proceedings of the 8th WSEAS International Conference on E-Activities and information security and privacy*. World Scientific, Engineering Academy, and Society (WSEAS). 2009, pp. 172–177.

[81] Steffen Pfrang et al. "Design and architecture of an industrial IT security lab". In: *International Conference on Testbeds and Research Infrastructures*. Springer. 2016, pp. 114–123.

[82] Cuong Pham et al. "CyRIS: A cyber range instantiation system for facilitating security training". In: *Proceedings of the Seventh Symposium on Information and Communication Technology*. ACM. 2016, pp. 251–258.

[83] Qais Qassim et al. "A Survey of SCADA Testbed Implementation Approaches". In: *Indian Journal of Science and Technology* 10.26 (2017). ISSN: 0974 -5645. URL: `http://www.indjst.org/index.php/indjst/article/view/116775`.

[84] Muhammad Azizur Rahman, Algirdas Pakštas, and Frank Zhigang Wang. "Network modelling and simulation tools". In: *Simulation Modelling Practice and Theory* 17.6 (2009), pp. 1011–1031.

[85] Theodore Reed, Kevin Nauer, and Austin Silva. "Instrumenting competition-based exercises to evaluate cyber defender situation awareness". In: *International Conference on Augmented Cognition*. Springer. 2013, pp. 80–89.

[86] Michael Richmond. "ViSe: A virtual security testbed". In: *University of California, Santa Barbara, Tech. Rep* (2005).

[87] Lee M Rossey et al. "Lariat: Lincoln adaptable real-time information assurance testbed". In: *Aerospace Conference Proceedings, 2002. IEEE*. Vol. 6. IEEE. 2002, pp. 6–6.

[88] Jose Rubio-Hernan, Juan Rodolfo-Mejias, and Joaquin Garcia-Alfaro. "Security of Cyber-Physical Systems". In: *Conference on Security of Industrial-Control-and Cyber-Physical Systems*. Springer. 2016, pp. 3–18.

[89] Julie A Rursch and Doug Jacobson. "This IS child's play Creating a "playground"(computer network testbed) for high school students to learn, practice, and compete in cyber defense competitions". In: *Frontiers in Education Conference, 2013 IEEE*. IEEE. 2013, pp. 1776–1778.

[90] Julie A Rursch and Doug Jacobson. "When a testbed does more than testing: The Internet-Scale Event Attack and Generation Environment (ISEAGE)-providing learning and synthesizing experiences for cyber security students." In: *Frontiers in Education Conference, 2013 IEEE*. IEEE. 2013, pp. 1267–1272.

[91] Wayne J Schepens et al. "The Cyber Defense Exercise: An evaluation of the effectiveness of information assurance education". In: *The Journal of Information Security* 1.2 (2002), pp. 1–14.

[92] Z Cliffe Schreuders et al. "Security Scenario Generator (SecGen): A Framework for Generating Randomly Vulnerable Rich-scenario VMs for Learning Computer Security and Hosting CTF Events". In: *USENIX*. USENIX Association. 2017.

[93] Rose Shumba. "Teaching hands-on Linux host computer security". In: *Journal on Educational Resources in Computing (JERIC)* 6.3 (2006), p. 5.

[94] C. Siaterlis and M. Masera. "A Review of Available Software for the Creation of Testbeds for Internet Security Research". In: *2009 First International Conference on Advances in System Simulation*. 2009, pp. 79–87. DOI: 10.1109/SIMUL.2009.33.

[95] Christos Siaterlis and Béla Genge. "Cyber-physical testbeds". In: *Communications of the ACM* 57.6 (2014), pp. 64–73.

[96] Christos Siaterlis and Marcelo Masera. "A survey of software tools for the creation of networked testbeds". In: *International Journal On Advances in Security* 3.2 (2010), pp. 1–12.

[97] Christos Siaterlis, Andres Perez-Garcia, and Marcelo Masera. "Using an Emulation Testbed for Operational Cyber Security Exercises". In: *International Conference on Critical Infrastructure Protection*. Springer. 2011, pp. 185–199.

[98] Shachar Siboni et al. "Advanced security testbed framework for wearable IoT devices". In: *ACM Transactions on Internet Technology (TOIT)* 16.4 (2016), p. 26.

[99] Austin Silva et al. "Factors impacting performance in competitive cyber exercises". In: *Proceedings of the Interservice/Interagency Training, Simulation and Education Conference, Orlando, FL*. 2014.

[100] Robin Snyder. "Ethical hacking and password cracking: a pattern for individualized security exercises". In: *Proceedings of the 3rd annual conference on Information security curriculum development*. ACM. 2006, pp. 13–18.

[101] Teodor Sommestad. "Experimentation on operational cyber security in CRATE". In: NATO STO-MP-IST-133 Specialist Meeting, Copenhagen, Denmark. 2015.

[102] Teodor Sommestad and Jonas Hallberg. "Cyber security exercises and competitions as a platform for cyber security experiments". In: *Nordic Conference on Secure IT Systems*. Springer. 2012, pp. 47–60.

[103] Yannis Soupionis and Thierry Benoist. "Cyber-physical testbed—The impact of cyber attacks and the human factor". In: *Internet Technology and Secured Transactions (ICITST), 2015 10th International Conference for*. IEEE. 2015, pp. 326–331.

[104] US Joint Staff. "Joint Training Manual for the Armed Forces of the United States (CJCSM 3500.03 D)". In: *Washington, DC: Joint Chiefs of Staff* (2012).

[105] Joseph Stites, Ambareen Siraj, and Eric L Brown. "Smart Grid Security Educational Training with ThunderCloud: A Virtual Security Test Bed". In: *Proceedings of the 2013 on InfoSecCD'13: Information Security Curriculum Development Conference*. ACM. 2013, p. 105.

[106] Georgiana Subaşu, Livia Roşu, and Ion Bădoi. "Modeling and simulation architecture for training in cyber defence education". In: *Electronics, Computers and Artificial Intelligence (ECAI), 2017 9th International Conference on*. IEEE. 2017, pp. 1–4.

[107] Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. "Cyber security of a power grid: State-of-the-art". In: *International Journal of Electrical Power & Energy Systems* 99 (2018), pp. 45–56.

[108] Pang-Wei Tsai and Chu-Sing Yang. "Testbed@ TWISC: A network security experiment platform". In: *International Journal of Communication Systems* 31.2 (2018), e3446.

[109] Pang-Wei Tsai et al. "Control frameworks in network emulation testbeds: A survey". In: *Journal of computational science* 22 (2017), pp. 148–161.

[110] Vincent Urias, Brian Van Leeuwen, and Bryan Richardson. "Supervisory Command and Data Acquisition (SCADA) system cyber security analysis using a live, virtual, and constructive (LVC) testbed". In: *Military Communications Conference, 2012-MILCOM 2012*. IEEE. 2012, pp. 1–8.

[111] Giovanni Vigna et al. "Ten Years of iCTF: The Good, The Bad, and The Ugly." In: *3GSE*. 2014.

[112] Alexander Volynkin and Victor Skormin. "Large-scale reconfigurable virtual testbed for information security experiments". In: *Testbeds and Research Infrastructure for the Development of Networks and Communities, 2007. TridentCom 2007. 3rd International Conference on*. IEEE. 2007, pp. 1–9.

[113] Jan Vykopal et al. "KYPO Cyber Range: Design and Use Cases". In: (2017).

[114] Jan Vykopal et al. "Lessons learned from complex hands-on defence exercises in a cyber range". In: *Frontiers in Education Conference (FIE)*. IEEE. 2017, pp. 1–8.

[115] Brian White et al. "An integrated experimental environment for distributed systems and networks". In: *ACM SIGOPS Operating Systems Review* 36.SI (2002), pp. 255–270.

[116] Gregory B White and Dwayne Williams. "The collegiate cyber defense competition". In: *Proceedings of the 9th Colloquium for Information Systems Security Education*. 2005.

[117] Christian Willems and Christoph Meinel. "Online assessment for hands-on cyber security training in a virtual lab". In: *Global Engineering Education Conference (EDUCON), 2012 IEEE*. IEEE. 2012, pp. 1–10.

[118] Christian Willems and Christoph Meinel. "Practical network security teaching in an online virtual laboratory". In: *Proceedings of the International Conference on Security and Management (SAM)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). 2011, p. 1.

[119] Bradley J Wood and Ruth A Duggan. "Red teaming of advanced information assurance concepts". In: *DARPA Information Survivability Conference and Exposition, 2000. DISCEX'00. Proceedings*. Vol. 2. IEEE. 2000, pp. 112–118.

[120] Evangelia Xypolytou et al. "The FUSE testbed: establishing a microgrid for smart grid security experiments". In: *e & i Elektrotechnik und Informationstechnik* 134.1 (2017), pp. 30–35.

[121] Shingo Yasuda et al. "Alfons: A Mimetic Network Environment Construction System". In: *International Conference on Testbeds and Research Infrastructures*. Springer. 2016, pp. 59–69.

A APPENDIX: Citation Data

No.	Paper Title	Citation Count	Year Published
1	Design of Cyber Warfare Testbed	2	2019
2	Cyber security of a power grid: State-of-the-art	10	2018
3	Testbed@ TWISC: A network security experiment platform	0	2018
4	Achieving reproducible network environments with INSALATA	1	2017
5	A Survey on Smart Grid Cyber-Physical System Testbeds.	52	2017
6	Capability Detection and Evaluation Metrics for Cyber Security lab Exercises	0	2017
7	Control frameworks in network emulation testbeds: A survey	1	2017
8	Cybersecurity training in control systems using real equipment	1	2017
9	Design and implementation of cybersecurity testbed for industrial IoT systems	2	2017
10	Developing a capability to classify technical skill levels within a Cyber Range	0	2017
11	Experiment as a service	1	2017
12	Extending Our Cyber-Range CYRAN with Social Engineering Capabilities	0	2017
13	Gamifying ICS security training and research: Design, implementation, and results of S3	2	2017
14	Improving and Measuring Learning Effectiveness at Cyber Defense Exercises	1	2017
15	Instrumentation Research Methods for Cyber Security	7	2017
16	KYPO Cyber Range: Design and Use Cases	9	2017
17	Lessons learned from complex hands-on defence exercises in a cyber range	2	2017
18	Modeling and simulation architecture for training in cyber defence education	0	2017
19	The FUSE testbed: establishing a microgrid for smart grid security experiments	1	2017

20	Towards a Unified Data Storage and Generic Visualizations in Cyber Ranges	1	2017
21	Using virtual environments for the assessment of cybersecurity issues in IoT scenarios	14	2017
22	Advanced security testbed framework for wearable IoT devices	18	2016
23	Alfons: A Mimetic Network Environment Construction System	4	2016
24	Cybervan: A cyber security virtual assured network testbed	11	2016
25	CyRIS: A cyber range instantiation system for facilitating security training	12	2016
26	Design and architecture of an industrial IT security lab	4	2016
27	Developing a distributed software defined networking testbed for IoT	7	2016
28	PowerCyber: A remotely accessible testbed for Cyber Physical security of the Smart Grid	5	2016
29	RIO: A denial of service experimentation platform in a Future Internet Testbed	0	2016
30	Security of Cyber-Physical Systems	1	2016
31	Softgrid: A software-based smart grid testbed for evaluating substation cybersecurity solutions	4	2016
32	Virtualization of industrial control system testbeds for cybersecurity	10	2016
33	A real-time testbed environment for cyber-physical security on the power grid	12	2015
34	Communications network analysis in a SCADA system testbed under cyber-attacks	9	2015
35	Cyber-attack and defense simulation framework	4	2015
36	Cyber modeling & simulation for cyber-range events	9	2015
37	Cyber-physical systems testbed based on cloud computing and software defined network	6	2015
38	Cyber-physical testbed—The impact of cyber attacks and the human factor	3	2015

39	Experimentation on operational cyber security in CRATE	2	2015
40	i-tee: A fully automated Cyber Defense Competition for Students	7	2015
41	KYPO–A Platform for Cyber Defence Exercises	7	2015
42	Sdsecurity: A software defined security experimental framework	44	2015
43	Understanding collaborative challenges in it security preparedness exercises	10	2015
44	Building a Virtual Cybersecurity Collaborative Learning Laboratory (VC-CLL)	1	2014
45	Cloud-based security research testbed: A DDoS use case	10	2014
46	Cloud-based testbed for simulation of cyber attacks	23	2014
47	Cyber-physical testbeds	17	2014
48	Cyber security backdrop: A scada testbed	13	2014
49	Factors impacting performance in competitive cyber exercises	19	2014
50	FITS: A flexible virtual network testbed architecture	42	2014
51	National cyber range overview	14	2014
52	Ten Years of iCTF: The Good, The Bad, and The Ugly.	30	2014
53	The design of ics testbed based on emulation, physical, and simulation (eps-ics testbed)	13	2014
54	A survey of cyber ranges and testbeds	15	2013
55	Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid	197	2013
56	Instrumenting competition-based exercises to evaluate cyber defender situation awareness	13	2013
57	SCADAVT-A framework for SCADA security testbed based on virtualization technology	25	2013
58	Smart Grid Security Educational Training with ThunderCloud: A Virtual Security Test Bed	1	2013
59	This IS Child's Play	2	2013

60	When a testbed does more than testing: The Internet-Scale Event Attack and Generation Environment (ISEAGE)-providing learning and synthesizing experiences for ...	5	2013
61	Amici: An assessment platform for multi-domain security experimentation on critical infrastructures	24	2012
62	Beyond network simulators: Fostering novel distributed applications and protocols through extendible design	6	2012
63	Cyber security exercises and competitions as a platform for cyber security experiments	20	2012
64	Cyber Security Assessment Tools and Methodologies	5	2012
65	Online assessment for hands-on cyber security training in a virtual lab	30	2012
66	Supervisory Command and Data Acquisition (SCADA) system cyber security analysis using a live, virtual, and constructive (LVC) testbed	36	2012
67	Towards an experimental testbed facility for cyber-physical security research	10	2012
68	A control system testbed to validate critical infrastructure protection concepts	69	2011
69	An overview of cyber attack and computer network operations simulation	28	2011
70	A testbed for analyzing security of SCADA control systems (TASSCS)	95	2011
71	DefEX: Hands-On Cyber Defense Exercise for Undergraduate Students	7	2011
72	Hit'em where it hurts: a live security exercise on cyber situational awareness	34	2011
73	Practical network security teaching in an online virtual laboratory	16	2011
74	Using an Emulation Testbed for Operational Cyber Security Exercises	2	2011
75	An experimental platform for assessing SCADA vulnerabilities and countermeasures in power plants	61	2010
76	An Intelligent network for federated testing of NetCentric systems	4	2010

77	A survey of software tools for the creation of networked testbeds	14	2010
78	Design of a virtual computer lab environment for hands-on information security exercises	13	2010
79	Organizing large scale hacking competitions	44	2010
80	The Blunderdome: An Offensive Exercise for Building Network, Systems, and Web Security Awareness.	6	2010
81	The DETER project: Advancing the science of cyber security experimentation and test	78	2010
82	Current developments in DETER cybersecurity testbed technology	26	2009
83	Guide for designing cyber security exercises	30	2009
84	Real-time security exercises on a realistic interdomain routing experiment platform	7	2009
85	The Cyber Scenario Modeling and Reporting Tool (CyberSMART)	3	2009
86	Network modelling and simulation tools	63	2009
87	Design on SCADA test-bed and security device	29	2008
88	Cyber attack modeling and simulation for network security analysis	82	2007
89	Large-scale reconfigurable virtual testbed for information security experiments	18	2007
90	The development and deployment of a multi-user, remote access virtualization system for networking, security, and system administration classes	84	2007
91	A virtual machine architecture for creating IT-security laboratories	13	2006
92	Ethical hacking and password cracking: a pattern for individualized security exercises	5	2006
93	Experience with deter: a testbed for security research	195	2006
94	Teaching hands-on Linux host computer security	3	2006

95	Exploring a national cybersecurity exercise for universities	78	2005
96	Rinse: The real-time immersive network simulation environment for network security exercises	82	2005
97	The INFN-grid testbed	9	2005
98	ViSe: A virtual security testbed	14	2005
99	An integrated experimental environment for distributed systems and networks	1667	2002
100	Lariat: Lincoln adaptable real-time information assurance testbed	112	2002

Table 13: Reviewed paper citation data as of Augusts 10 2018

Muhammad Mudassar Yamin

Mudassar is currently a PhD candidate at Norwegian University of Science and technology, his focus of research is systems security. He holds multiple cyber security certifications like OSCP,LPT-MASTER,CEH,CHFI. A list of his publication can be found here:

https://scholar.google.no/citations?user=Do_xVQMAAAAJ&hl=en

Basel Katt

Basel is currently working as an Associate Professor in Norwegian University of Science and technology. His focus of research is:

- Software security and security testing
- Software vulnerability analysis
- Model driven software development and model driven security
- Access control, usage control and privacy protection
- Security monitoring, policies, languages, models and enforcement

A list of his publication can be found here:
https://wo.cristin.no/as/WebObjects/cristin.woa/wa/fres?sort=ar&pnr=811108&action=sok

VASILEIOS GKIOULOS

VASILEIOS is currently working as a post-doctoral research fellow. His are of research is security of cyber physical systems. A list of his publication can be found here:

https://scholar.google.no/citations?user=Jgo7_q4AAAAJ&hl=en