EDF-2021-CYBER-D: Improved capacity for cyber training and exercises

Proposals are invited against any of the following topic:

EDF-2021-CYBER-D-IECTE: Improved efficiency of cyber trainings and exercises

Budget

The Union is considering a contribution of up to EUR 20 000 000 to support proposals addressing any of the abovementioned topic and its associated specific challenge, scope, targeted activities and main functional requirements.

Up to one action may be funded under this call.

The Commission will pay particular attention to the other R&D and dual-use on-going initiatives at Union level to avoid unnecessary duplication.

Specific challenge

Personnel development is one of the key requirements for effective cyber defence. Extensive trainings and exercises constitute the best means to enhance and validate the skills of the cyber defence workforce. For this, Member States have invested in establishing cyber ranges that provide controlled artificial environments where, among others, malicious activities can be simulated without negative impact on live systems in an organization. However, the existing cyber ranges can be developed further to achieve their full personnel development potential. In turn, it supports cyber operators improving their skillset and benefits military commanders in understanding cyber as a cross-domain challenge. This includes addressing threats and opportunities driven from the emerging disruptive technologies.

The proposals are expected to address following challenges:

First, the maturity level of user simulation running on cyber ranges is low. The user simulation is often limited to traffic generators and tools for testing user interfaces for a well- defined purpose. User simulation, which leaves a non-detectable footprint and produces logs while being indistinguishable from the real human users', is needed for providing more meaningful, realistic and life-like scenarios for exercises and training sessions. Additionally, scenarios that rely on the actions of simulated users, e.g. because of phishing emails, require a solution that gives the training or exercise organizer control over the user simulation to make sure simulated users act in accordance to the scenario.

Second, cyber ranges lack capabilities to assess and reflect the decision-making process of cyber operators during an exercise or training. Current systems fail to provide insight to the particular actions cyber operators perform to achieve the objectives. This includes unanswered questions such as which tools and commands were used and options selected in the graphical interface, who was communicated with (both online and in the war-room), what was searched online (and whether that was useful). This presents additional challenges as systematic monitoring and assessment of skill gaps is not possible, especially for more complex exercises. However, automated performance assessment and analysis of the participants allows the training and exercise instructors to monitor either on individual or team level their performance in more detail.

Third, scenarios involving user-simulation and systems enabling analysis and assessment of the decision-making process of cyber operators should be accessible and interoperable for different cyber ranges. This can be enabled through scenario development language. Besides potential cost-efficiency, it also improves and upgrades the scenarios over time based on feedback by many users. Hence, simulated users, scoring system and the analysis of the operators' performance can be accordingly elevated, training and exercises therefore continuously improved.

Fourth, so far, many cyber ranges focus only on one domain and its functionalities, but the impact of cyber-attacks must be considered as a cross-domain challenge. Therefore, a multi-domain cyber range simulation must support and simulate land, air, sea and space domains. This includes, for example, military systems (e.g. battle management systems), radio and operational technology. Especially the integration of the electromagnetic spectrum (EMS) and the common understanding of cyber and EMS should be a key factor of future cyber ranges. The challenge is to support a highly realistic simulation of multiple domains, the interconnection of systems and to assess the impact of the inter-dependencies between those systems. Such simulations would support the training and evaluation of multi-domain common operating pictures and its operations as well as development and testing of new military approaches and doctrines to cyber/EMS threats.

Last, conducting large-scale cyber exercises or simulating real-life modern ICT environment requires unique and complex set of capabilities and infrastructure (such as specialized hardware to simulate cellular networks, industrial controllers and other parts of critical infrastructure). The most practical way to create such complex environments is through cooperation among Member States and federating cyber range infrastructure and exercise content. Such approach requires development of common standards, protocols, and software solutions to allow federated scoring and situational awareness throughout the federated environment.

Scope

The objective of this topic is to create a toolset that allows significantly increased efficiency in the cyber trainings and exercises process while also enhancing cyber ranges interoperability and cost-efficiency, taking into account challenges described.

To develop a technological demonstrator modules that can be easily configured and interfaced to existing system used to conduct cyber trainings and exercises. Integration of the technologies must be demonstrated within TRL 4-8, but specific TRL may wary depending on the work package.

Simulated users

Development of agents capable of using common software applications in a similar manner to human users. Actions in scope for the user simulation is benign use of common software applications (e.g. word processors, web browsers, file management applications, and email clients). The solution is able to replay sequences such actions obtained from the automated performance analysis when human operators has used systems and allows custom behaviours to sequences of such actions to be crafted. Furthermore, the solution also allows custom modifications of action sequences obtained from the automated performance analysis, e.g. to modify a sequence of actions to create an alternative scenario. Simulated users be used to, among other things, simulate social engineering incidents (e.g. phishing), the generation of

system logs, and the generation of network logs. The footprint of the simulated users is indistinguishable from those of real users in logs and must be visible in graphical user interfaces.

Automated performance analysis

Development of a system capable of collecting data about the cyber operators activities during trainings and exercises, and automatically analysing tasks of low to medium complexity while also providing supporting data and insight to evaluate advanced situations which are often difficult to fully assess through automated techniques. This should be based on combining already well-researched and documented methods (e.g. application programming interfaces (APIs)), and big data analysis, image analysis and neural language processing to collect and analyse cyber operator's behaviour during training and an exercise.

Scenario development language

Scenario development language enables scenario sharing with other cyber ranges and ensures interoperability between different cyber ranges. While the cyber ranges are interoperable regarding more common activities (training, exercises etc.) and more complex scenarios can be implemented raising the overall preparedness of cyber operators utilizing the capabilities. The language itself should be described based on research of existing cloud and virtualization topologies and should be extended with specific components (simulation, scoring, federation etc. attributes) that are needed in the cyber range environments.

Multi-domain simulations

Development of enhanced multi-domain cyber range simulations for at least 2 domains (e.g., land, sea and/or space) and the standards and interfaces to interconnect relevant systems and environments (e.g., battle management systems, EMS or other systems) in order to allow simulation of realistic joint cross-domain scenarios and situational awareness.

Situational awareness and scoring

The activities include developing standards and protocols for federated scoring system, exchange of situational awareness information, including federated cyber range operation.

Targeted activities

The proposals must cover the following activities as referred in article 10.3 of the EDF Regulation, not excluding upstream or downstream activities eligible for development actions if deemed useful to reach the objectives:

- Studies, such as feasibility studies to explore the feasibility of new or improved technologies, products, processes, services and solutions;
- The design of a defence product, tangible or intangible component or technology as well as the definition of the technical specifications on which such design has been developed which may include partial tests for risk reduction in an industrial or representative environment;

- The development of a model of a defence product, tangible or intangible component or technology, which can demonstrate the element's performance in an operational environment (system prototype);
- The testing of product, tangible or intangible component or technology.

The proposals must address in particular the following:

Simulated users

- Study the methods and technologies to develop simulated user;
- Develop a prototype for simulated user capable of using common software applications while producing realistic logs and footprints in the machines;
- Develop a method for converting recorded behavioural data of cyber operators that can be used in the user simulation;
- Proof of concept testing and validation of the proposed toolset to present that the user simulation produces a footprint similar to normal users when it is applied.

Automated performance analysis

- Study the methods and technologies that provide information about the decision-making process during a training or exercises (i.e. consoles, graphical interfaces (including videos from environment), network traffic, audio between team members));
- Study the technologies to gather, store and process information produced during a training or exercise;
- Design methodology to correlate cyber-attacks data and collected cyber operators behavioural data, and the ways to improve the individual and collective performance of cyber operators. The methodology should also provide feedback loop for exercise designers to improve the learning effect of the exercises;
- Design methods to ensure the integrity of the automated analysis process and/or consider classification of military environments;
- Develop a prototype for the automated analysis that includes at least logging, data parsing and performance evaluation functionalities;
- Proof of concept testing and validation of the proposed toolset. Scenario development language
- Study existing cloud and virtualization topologies, including defining data format to define a common scenario language that can be used by different cyber ranges;
- Develop an extendable scenario development language in coherence with the other capabilities described in this call;
- Proof of concept testing and validation of the scenario development language on an existing cyber range.

Multi-domain simulations

- Develop standards and interfaces for interconnecting multi-domain cyber ranges (e.g., land, air, sea and space) for cyber trainings and exercises;
- Develop multi-domain scenarios for capacity building;
- Proof of concept testing and validation of the proposed toolset.

Situational awareness and scoring

- Study the existing technology solutions used by Member States for cyber ranges' situational awareness and implement the solution in federated environment;
- Design common standards for cyber ranges' situational awareness.

Functional requirements

General

- The methods and technologies proposed and developed can be implemented in most existing cyber range environments.

Simulated users

- The simulated users must be able to perform common actions in commonly used enterprise software (e.g. word processors, web browsers, file management applications, and email clients);
- The simulated users must produce logs in the machines that, using commonly used log analysis tools, cannot be distinguished from the logs produced by humans performing the same actions;
- The actions performed by the simulated users must be possible to schedule using recorded actions of human operators, with a crafted specification detailing actions, and through the combination of recorded actions and crafted specifications.

Automated performance analysis

- Cyber operators in the training or exercise environment are identified and the information processed is linked to specific operators, which allows monitoring and automatic analysis of cyber operators' activity on individual and team level;
- The analysis of the cyber operators' behaviour must be done in near-real time, and there must be an option to customize the set of-rules and parameters depending on the training or exercise;
- Reporting of the analysis of the cyber operators' activity should be displayed according to the timeline of the exercise or training;
- Both cloud computing and on-premise solutions must be available for the cyber range operators depending on the classification level of the environment.

Scenario development language

- Scenario development language should take into account previously conducted related research and development, i.e. should have a functionality similar to OASIS Topology as well as Orchestration Specification for Cloud Applications (TOSCA) version 1.3[43], OpenStack or other cloud architecture solutions;
- Scenario development language must be machine and human-readable, use English syntax and be based on textual language (also have graphical representation), be scalable and extendable, enable using regular expressions and analysing capabilities, enable commenting the content;

- Scenario development language should at minimum support definitions such as network, virtual machines, external component configuration, and post-deployment actions;
- Scenario development language should allow post-deployment actions, e.g. verification of the deployed system against the planned description;
- The language for post-deployment actions (e.g. simulated users and attacks) should provide control over the timing of events and enable the creation of dynamic plans where actions depend on the current state of the cyber range environment and the outcome of previous actions;
- Scenario development language should enable a common analysis of indicators of compromise (IOC) and malicious artefacts (operational level);

Multi-domain simulations

- The proposed solutions on cyber range should provide simulations for at least two domains;
- The simulations should not only include simulations on the network layer and above (Open Systems Interconnection, (OSI) model layer 3+), but also include data link and physical layers (OSI model layer 1-2).

Situational awareness

- Situational awareness protocols and solution must be sufficient for tracking typical large scale cyber exercise needs[1];
- Existing software solutions must be taken into account when developing standards, protocols, and solutions to ease integration and interoperability.

Expected impact

- Advanced preparedness of cyber defence operators and the capacity and interoperability of cyber ranges, which contributes cost-efficiently to EU cyber defence posture.
- The methods and tools provided will offer a better understanding of the decision-making process, both on the individual and team levels. Individuals and teams can have *post mortums* to gather lessons identified from the training or exercise. The results must also be comparable between several trainings and exercises. A strong data-based approach supports organizers in performing analytic evaluation to improve the quality of an exercise, evaluate its impact on learners compare the effectiveness of different approaches, and refine and enrich training scenarios.
- Based on the methods and tools developed for understanding the operator's decision-making process, a framework for modelling simulated users can be developed. This capability reduces the need for human interventions but increases the quality of the cyber trainings and exercises. Realistic normal system usage also serves as background traffic in intrusion detection tests, thereby supporting the development of tools for cyber situational awareness.
- The ability to facilitate cyber trainings and exercises in a federated environment will increase cooperation among member states, efficiency of organizing large-scale

[1] A large scale cyber exercise can be understood for instance as an exercise that has more than 100 participants and 50 observers.

exercises, and complexity that can be achieved realistically in such exercises. Thus, improving the overall effectiveness of cyber trainings.

- These benefits will ultimately result in lower cost and larger number of successful trainings. Better-organized exercises will help to gain skill improvements quicker and with less hours of training.