Automated Cyber Red Teaming

Paloma de la Vallée Georgi Nikolov Thibault Debatty Wim Mees [1]

[1]Royal Military Academy, Brussels, Belgium

Objective

Our corporate information networks and systems are being actively targeted by threat sources on a daily basis. Detecting incidents and responding to them requires a very specific set of knowledge, skills, and abilities (KSA). Developing the cyber-related KSA that are required for each role in the organization typically requires hands-on training on a cyber range. In order to develop adequate mental models, the training scenarios that are used must be sufficiently realistic and therefore inevitably end up being complex to implement and execute.

Running a training scenario on a cyber range therefore currently requires cyber experts to play the role of the attackers. These are referred to as the "red team". Given the limited number of cyber experts that are available, unfortunately this solution does not scale very well.

This project aims at developing a system able to take over the role of the red team, executing a pre-designed attack independently on a network. This system has the capability to orchestrate a series of action in a timely manner.

This will make it possible to organize an ambitious hands-on cyber training program without creating an excessive load on the expert cyber training staff. Finally, not only will automated red teaming make it possible to organize larger numbers of training sessions, but it will also result in repeatable red team performance, which is important for evaluating and possibly certifying skills.

Architecture

At present, the typical use case is the automation of a predefined attack to be carried out on a scenario deployed on a cyber range.

The design of the attack is performed by a cyber expert, responsible to develop the network scenario and the associated suite of actions to complete the attack.

The automation of the red team attack is assured by an orchestrator centralising the decision process, running on a specific machine comprised in the network scenario. It sends instructions to attacker machines, typically Kali virtual machines (VMs), and retrieves the outcome of the actions for further processing. It is responsible for launching the actions at the appropriate time.

From the orchestrator perspective, the attack is organised as an oriented and conditional graph of tasks. The children tasks are launched depending on the status of the parents, i.e. whether the parent tasks failed, succeeded, or simply have been evaluated. The graph allows to define parallel branches. Depending on the user requests, the different branches can run in parallel; or the orchestrator will switch from one branch to the next, should a preceding branch fail. The orchestrator is framework-agnostic and can execute all instructions that are managed by command line, both on Linux-based and Windows machines.

Instance of automated attack

The orchestrator has already been successfully used to automate different attacks.

One of these instances is presented here.

The network scenario depicted on Figure 1 is deployed on the Cylab cyber range [1].

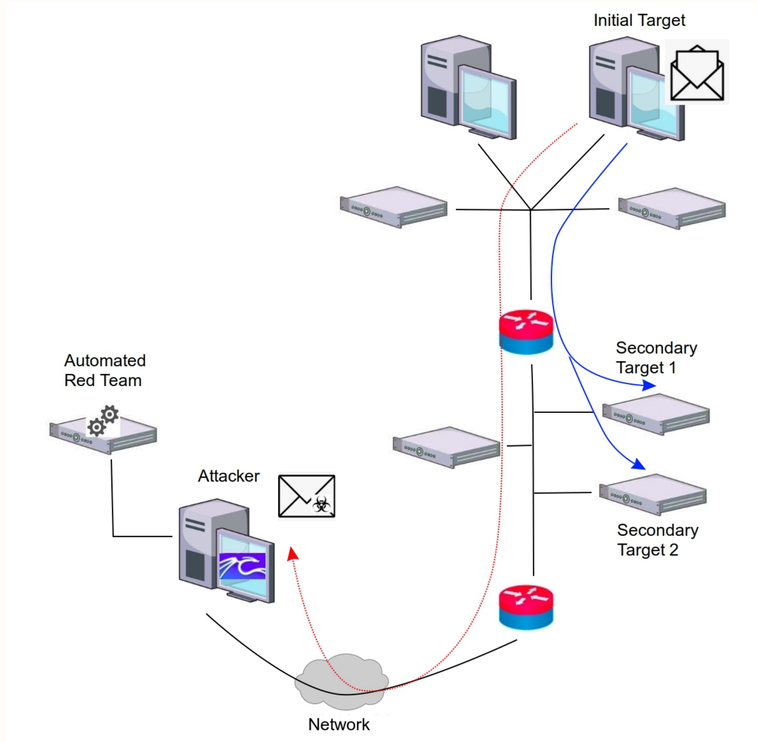

Figure 1: Network and attack scenario

The red team is represented on the left, and consists of a Kali VM, or Attacker, and the orchestrator or Automated Red Team.

The target organisation network is represented on the right. It consists of two LANs, interconnected via routers also having the function of firewalls. The lower LAN has different Linux servers holding data, while the upper LAN has some workstations and some servers. The firewalls only allow outgoing traffic, except for established connections. Both sides can connect via a network mimicking the Internet. However, due to the firewalls, the red team cannot directly scan or access the machines in the target organisation and must manage to get initial access and a foothold by some other means.

In this scenario, the attack is designed as follows:

- On the Attacker VM, a specific payload is created with help of the Veil framework [2].
- The payload is sent by mail in a password-protected encrypted attachment to the user connected on the Initial Target.
- Opening the email and running the attachement opens a connection towards the Attacker from behind the firewalls (red dotted line), resulting in a meterpreter session.
- Once the Initial Target is compromised, the Attacker starts by migrating the meterpreter session to a stable process and elevating the privileges.
- The Attacker then explores the machine and acquires two separate sets of credentials valid on Secondary Targets.
- From the Initial Target used as pivot, the attacker connects to the Secondary Targets (blue lines) to download sensitive material.

The graph of tasks associated with this attack is depicted on Figure 2.

In this attack, all children tasks await the success of their parent task(s) before being launched. The graph shows the different tasks organised in three branches that are processed in parallel with dependencies between the branches: the mail will only be opened (branch 3) once the meterpreter listener is launched (branch 1) and the mail has been sent (branch 2).

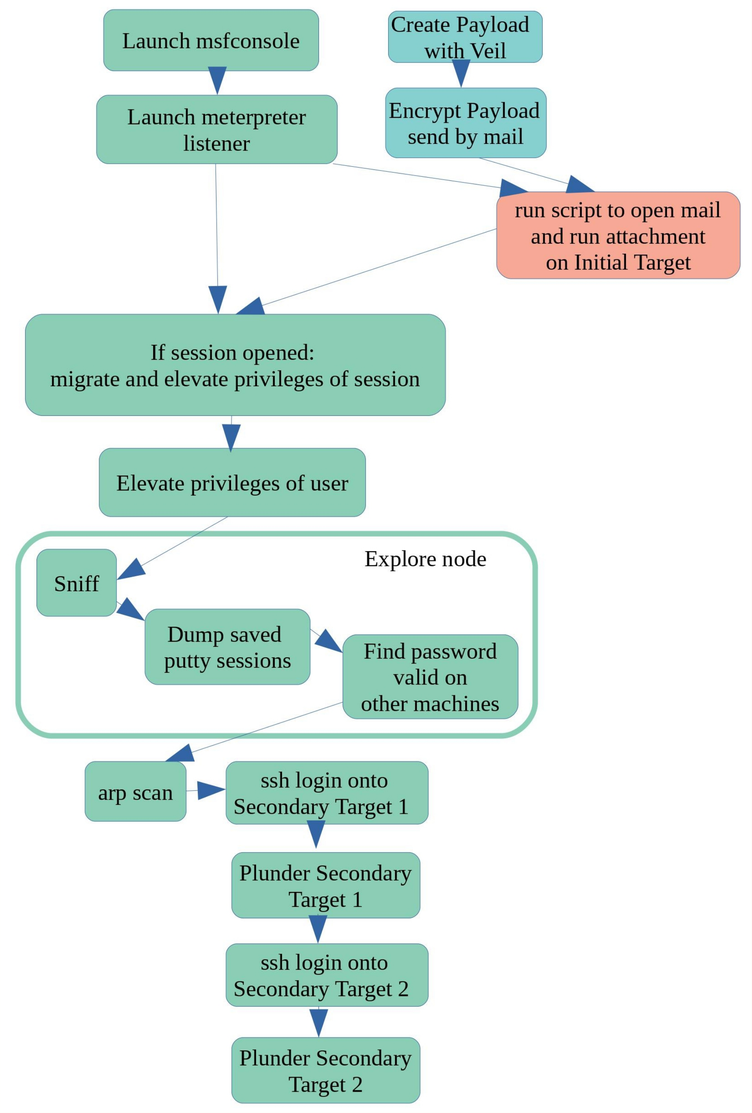

Figure 2: Attack graph

This attack demonstrates the ability of the orchestrator to manage complex tasks inter-dependencies in a timely manner. It also proves the integration of the orchestrator with different frameworks (Veil, metasploit), as well as the successful execution of instructions on Linux-based machines (Kali, Secondary Targets) and Windows machine (Initial Target).

Future work

In its present state, the automated red team executes a pre-defined series of actions designed by a cyber-expert for a specific scenario, with predictable outcomes. In the future, an attack planner could be developed. Provided with a network configuration in an appropriate form, this planner could be able to find an attack path to reach a certain goal, such as compromising a specific machine in the network [3]. Further down the line, the planner might possibly be replaced by an Artificial Intelligence, able to sense its way into the network and take appropriate action based on a limited view of the network state and configuration.

Acknowledgements: This work was supported by the ECHO project which has received funding from the European Union's Horizon 2020 research and innovation program under the grant agreement no 830943.

References

[1] T. Debatty, W. Mees, Building a Cyber Range for training CyberDefense Situation Awareness,2019 InternationalConference on Military Communications and Information

[2] https://github.com/Veil-Framework/Veil

[3] J. Yuen, Automated Cyber Red Teaming, 41 p, 2015. Available online: http://www.dtic.mil/dtic/tr/fulltext/u2/a618584.pdf