Reinforcing Cybersecurity Hands-on Training With Adaptive Learning

Pavel Seda, Jan Vykopal, Valdemar Švábenský, Pavel Čeleda
seda@fi.muni.cz

Masaryk University, Czech Republic

October 2021 @ FIE'21 conference

```
ACCOUNT CHECK: [ssh] Host: 10.10.10.4 (1 of 1, 0 complete) User: bob (1 of 1, 1 complet
e) Password: scorpio (36 of 100 complete)
ACCOUNT CHECK: [ssh] Host: 10.10.10.4 (1 of 1, 0 complete) User: bob (1 of 1, 1 complet
e) Password: qwerty (37 of 100 complete)
ACCOUNT CHECK: [ssh] Host: 10.10.10.4 (1 of 1, 0 complete) User: bob (1 of 1, 1 complet
e) Password: cameron (38 of 100 complete)
ACCOUNT CHECK: [ssh] Host: 10.10.10.4 (1 of 1, 0 complete) User: bob (1 of 1, 1 complet
e) Password: tomcat (39 of 100 complete)
ACCOUNT CHECK: [ssh] Host: 10.10.10.4 (1 of 1, 0 complete) User: bob (1 of 1, 1 complet
e) Password: testing (40 of 100 complete)
ACCOUNT CHECK: [ssh] Host: 10.10.10.4 (1 of 1, 0 complete) User: bob (1 of 1, 1 complet
e) Password: mountain (41 of 100 complete)
root@kali:~# ssh bob@10.10.10.4
bob@10.10.10.4's password:

The programs included with the De
the exact distribution terms for
individual files in /usr/share/do

Debian GNU/Linux comes with ABSOL
permitted by applicable law.
Last login: Mon Aug 20 10:13:40 2
bob@server:~$
```

```
individual files in /usr/share/doc/*/copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Thu Aug 23 11:39:31 2018 from 10.10.10.2
bob@server:~$ cd /tmp && make
g++ -Wall -pedantic -O2 -std=c++11 -pthread -o dcow dcow.cpp -lutil
bob@server:/tmp$ .dcow -s
-bash: .dcow: command not found
bob@server:/tmp$ ./dcow -s
Running ...
Password overridden to: dirtyCowFun

Received su prompt (Password: )

root@server:~# echo 0 > /proc/sys/vm/dirty_writeback_centisecs
root@server:~# \cp /tmp/.ssh_bak /etc/passwd
root@server:~# \rm /tmp/.ssh_bak
root@server:~# useradd innocentUser
root@server:~# passwd innocentUser
Enter new UNIX password:
Retype new UNIX password:
passwd: password updated successfully
root@server:~# usermod -aG sudo innocentUser
```

Problem Statement

Training input constraints

- High diversity of participants
- Different types of events
  - Arbitrary participants (students or professionals) for the same training instance

Training output implications

- High failure rate (around 50%)
  - Reduced training experience
  - Reduced learning outcomes

Goal of the Paper

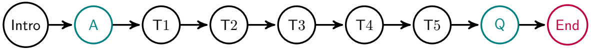
Design a training format and model that assigns suitable tasks for each participant based on their knowledge and skills.

Expectations

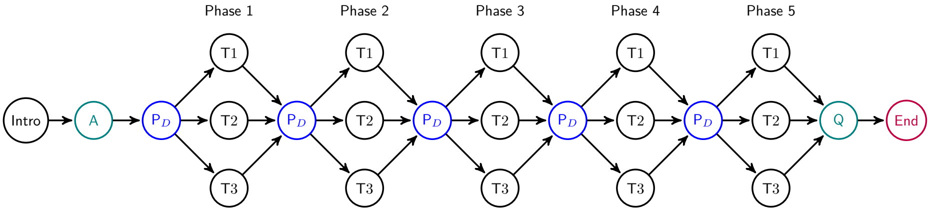
- Increased learning efficiency
- Increased learning experience
- Decreased training failure rate
- The same training can be used for a wider audience
- Participants finish the training in an allocated time

Training Format Design

Current format



Proposed format



- A is pre-training assessment, T_x is a task x, Q is a post-training questionnaire, and P_D is a phase decision node

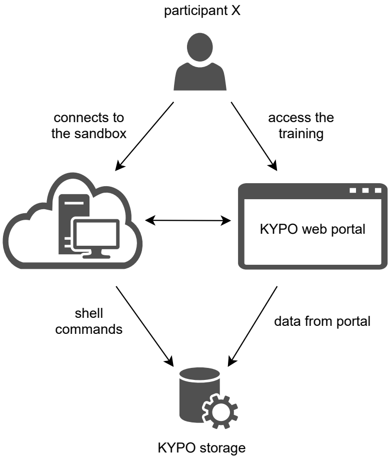
Model Design – Collected Metrics From Participants

Data from portal

- Pre-training assessment
- Submitted answers
- Task completed time
- Solution displayed

Data from sandboxes

- Shell commands



Model Design – Model

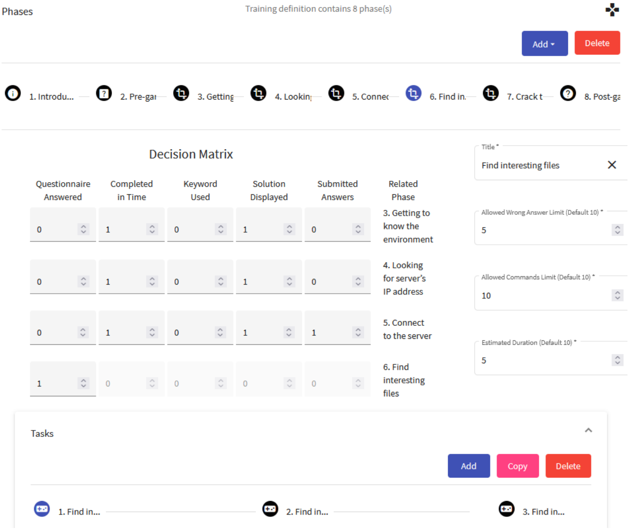
- The model uses defined metrics to evaluate the participants' performance and to assign a suitable task

$$\boldsymbol{W}^{(x)} = \left(w_{ij}^{(x)}\right), i = 1, \ldots, m, \quad j = \alpha, \beta, \gamma, \delta, \varepsilon \tag{1}$$

$$f(x) = \frac{\sum_{i=1}^{x} \left[p_i w_{i\alpha}^{(x)} + s_i \left(k_i w_{i\beta}^{(x)} + a_i w_{i\gamma}^{(x)} + t_i w_{i\delta}^{(x)} + w_{i\varepsilon}^{(x)}\right)\right]}{\sum_{i=1}^{x} \left(w_{i\alpha}^{(x)} + w_{i\beta}^{(x)} + w_{i\gamma}^{(x)} + w_{i\delta}^{(x)} + w_{i\varepsilon}^{(x)}\right)} \tag{2}$$

$$T_x = \begin{cases} n_x, & \text{if } f(x) \text{ is equal to } 0 \\ \operatorname{trunc}(n_x[1 - f(x)]) + 1, & \text{otherwise} \end{cases} \tag{3}$$

Setting Up the Model in the KYPO Learning Platform



Phases

Training definition contains 8 phase(s)

1. Introdu... — 2. Pre-gar — 3. Getting — 4. Lookin; — 5. Connec — 6. Find in. — 7. Crack t — 8. Post-ga

Decision Matrix

Questionnaire Answered	Completed in Time	Keyword Used	Solution Displayed	Submitted Answers	Related Phase
0	1	0	1	0	3. Getting to know the environment
0	1	0	1	0	4. Looking for server's IP address
0	1	0	1	1	5. Connect to the server
1	0	0	0	0	6. Find interesting files

Title *: Find interesting files

Allowed Wrong Answer Limit (Default 10) *: 5

Allowed Commands Limit (Default 10) *: 10

Estimated Duration (Default 10) *: 5

Tasks

1. Find in... — 2. Find in... — 3. Find in...

Model Limitations

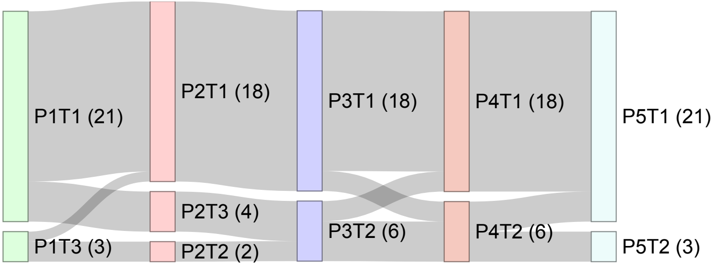
- The students' performance in a phase is evaluated in the same way in all tasks
- The observed metrics are binary
- The participants go through the training phase by phase
- It relies on the defined metrics (however, it can be enhanced or modified easily)

Model Evaluation – Case Study

- **Context and participants**
  - 24 participants (split among three events)
  - University and professional learners
- **Learning environment**
  - KYPO Cyber Range Platform
  - See our FIE'21 paper *Scalable Learning Environments for Teaching Cybersecurity Hands-on* `https://muni.cz/en/research/publications/1783808`
- **Adaptive training instance**
  - Linux tools, port scanning, secure shell, secure copy, and cracking ZIP files

Case Study Results

- Results from post-training questionnaire:
  - Participants reported that the training difficulty was adequate
  - 88% of the participants finished the training without taking a solution
- Participants’ transitions through the training



Recommendations for Instructors

- The pre-training assessment questionnaire should be simple and brief
- Adjust the weights in the model carefully
- Design at least three tasks for each phase
- Allocate more time for participants to complete the base phases than you expect

Conclusion

Traditional approach

- Difficult to accomplish training outcomes for a wider audience
- High failure rate

Research to practice – adaptive training instances

- Proposed model for cybersecurity adaptive training
- Improved participants' experience
- Decreased training failure rate
- Training instances can attract wider audience

Ongoing work

- Verification of the model with a larger amount of training instances and events

Stay in Touch

Get notified about the upcoming follow-up papers on the adaptiveness of cybersecurity training.

`https://twitter.com/cybersecmuni`

Thank you! Questions and feedback are welcome.

You can also e-mail me at `seda@fi.muni.cz`.

MUNI

FACULTY

OF INFORMATICS