On the design of security games: From frustrating to engaging learning

2016 USENIX Advances in Security Education Education Workshop
August 9, 2016

Jan Vykopal, Miloš Barták
Masaryk University, Brno
Who am I?

- Ph.D. graduate in flow-based intrusion detection.
- Founder and head of a university CSIRT in the Czech Republic.
- Researcher with KYPO – academic cloud-based cyber range.
- Coordinator and designer of hands-on training at KYPO platform, e.g. Czech national cyber defence exercise.
Outline

- KYPO game
  - Generic module of KYPO cyber range for running CtF games
  - Prototype game
  - Lessons learned
- Extensions of KYPO game
- Research questions
- User study – setup and results
- Conclusion and future work
KYPO cyber range

On the design of security games: From frustrating to engaging learning
Jan Vykopal, Masaryk University
KYPO game - design

- One educational use case of KYPO cyber range, implemented as a portlet.
- Framework for creating and running *attack-only* capture-the-flag games.

- Each game is split to several levels, players search for correct answer (flag).
- Each level offers hints that can be displayed in exchange for penalty points.
KYPO game – prototype

- Prototype game for teaching penetration testing.
- Four levels with the ultimate objective of NTP DoS amplification attack.
- Each player has own sandbox with a machine under control.
KYPO game – extensions I

Lessons learned:
- Difficulty of levels is not balanced.
- Learners are hesitant whether the hints will help them.
- Game-related information provided outside the platform are inconvenient.

Extensions:
- Improved hint system
  - Hints about hints available
    what tool to use, how to use the tool, ...
  - Players can now choose hints in arbitrary order.
- Embedded level solutions
  - Step-by-step tutorial for each level.
Lesson learned:
- Teacher has no information about the learners’ performance and progress in the ongoing event.

Extension:
- **Logging the learner’s actions**
  - Generic approach independent on specific game and its sandbox (hosts, network connections).
  - Captures only the interaction of the player and KYPO portal.
  - Does not capture any events or states from sandbox.
Research questions

1. **How helpful are the hints and solutions for the learners?**
   How do they contribute to completion of the level?

2. **What can be predicted from the participants’ actions?**
   What do game logs tell about the game and progress of the players?
Evaluation of extensions – setup I

- Experiment with a new game using the new features
  - More levels
  - Used improved hint system
  - Level solutions available.

- 21 participants in total - a diverse mix of players
  - Various level of experience and work positions (students, IT staff, researchers, experts).
  - Various European nations.
  - Various experience with hands-on training in cyber security.
Evaluation of extensions – setup II

- Self-assessment questionnaires for players
  - before the game,
  - after each level,
  - after the game.

  - *How was the level difficult?*
  - *Were the hints helpful?*
  - *Was the time limit sufficient?*
  - *What have you learned?*
  - *Would you like to play another game?*

- Game events of all players logged – a complement to self-assessment data.
Evaluation of extensions – hints

<table>
<thead>
<tr>
<th>Level</th>
<th>Learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X X</td>
</tr>
<tr>
<td>4</td>
<td>X X</td>
</tr>
<tr>
<td>5</td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>X X X X X</td>
</tr>
</tbody>
</table>

- New hint system used in 28 % of cases (arbitrary order of hints = green boxes).
- 77 % of all levels where learners opted for a hint(s) were then accomplished.
- **Mismatch of game logs and self-assessment** (double checked).
Evaluation of extensions – level solutions

- If the hints do still not help, and participants cannot proceed further, they can access the solution of the level.
- Contribution of solutions to accomplishment of the level was weaker than expected
  - Solution displayed and then the correct flag submitted – 60 % (17x)
  - Solution displayed and then the level skipped – 40 % (11x)
- Some participants might be frustrated and just wanted to enter the next level(s).
Evaluation of new features – game logs

- Time limit
- Median
- Min
- Max

On the design of security games: From frustrating to engaging learning
Jan Vykopal, Masaryk University
Evaluation of new features – game logs

On the design of security games: From frustrating to engaging learning
Jan Vykopal, Masaryk University
Conclusions

- Logging the game’s events provide useful data for analysis of game sessions to make them more engaging and fun.
- It is also useful for teachers to monitor ongoing session.

- Learners did use redesigned hint system and recommended solutions.
  - Evidence found in collected game events and the supplemental user survey.
  - Learners’ answers neither confirm nor disprove the benefit of the hints and solutions used.

- Other games events matched the learners’ assessment (level difficulty and duration).

- Future work: Do user surveys represent reliable tools for designing and evaluating hands-on training?
QUESTIONS?

THANKS FOR YOUR ATTENTION!

www.kypo.cz

@csirtmu

Jan Vykopal

vykopal@ics.muni.cz