MUNI
C4E

Benefits and Pitfalls of Using Capture the Flag Games in University Courses

Jan Vykopal, Valdemar Švábenský, Ee-Chien Chang
vykopal@ics.muni.cz

A part of this work was done during a stay of the first author at National University of Singapore.

SIGCSE'20, Portland, Oregon, USA

Capture the Flag game (CTF)

- Teams of players solve cybersecurity problems of varying complexity in a limited time (hours to days)
- Engaging education tool
- Increasingly popular in recent years
- Most CTFs focused on competition among teams

Cybersecurity university courses

- Feature lectures, tutorials, hands-on lab sessions, and homework assignments
- Taken by individuals who are graded based on summative assessment

Goal of this work

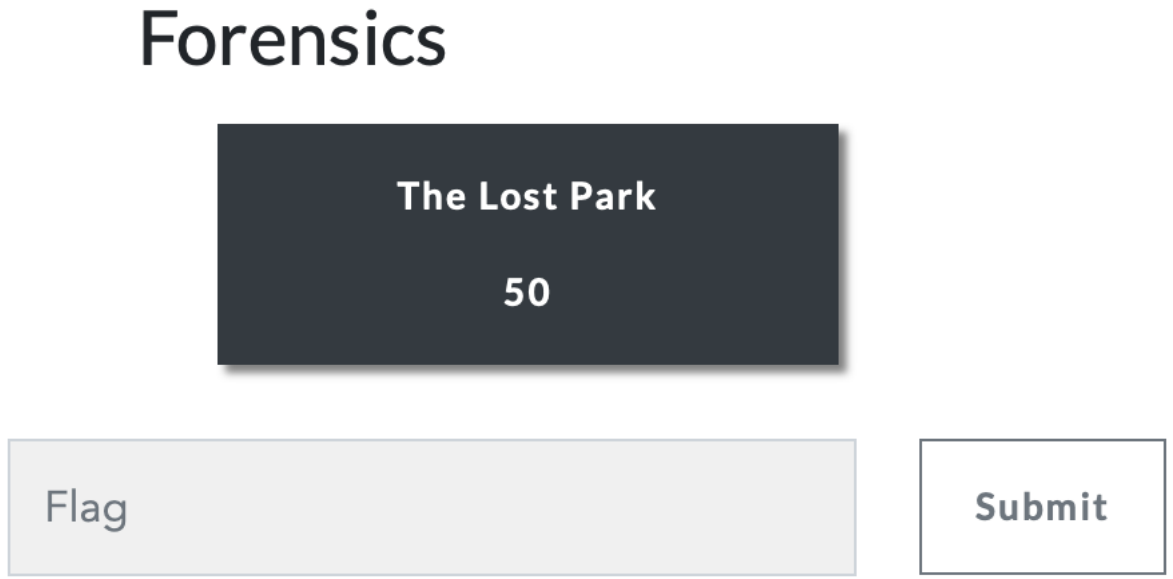
- Summarize experience from using CTF games as homework assignments in an introductory undergraduate course
- We identified four important aspects affecting the design of CTFs:
  - scoring,
  - scaffolding,
  - plagiarism,
  - learning analytics capabilities of the CTF platform.

Teaching context

- Introductory course on information and system security
- Taught in English at National University of Singapore in 2018/2019
- 37 students agreed to participate in the study (out of 120)
- Only two participants had played any CTF game before

CTF games features

- Challenge value
- Immediate response
- Scoreboard
- Hints for free (scaffolding)
- Challenge chains (dependencies)
- Run at CTFd (ctfd.io)



MUNI C4E

Collected data

- Game events generated by the students from the web-based CTF portal (flag submissions, hints taken)
- Answers from two surveys (pre, post)
- Students' marks from other forms of summative assessment of the course (midterm, final exam)

CTF content and parameters

- Two homework assignments – A1 and A2
- Various difficulty indicated by a category (basic, medium, advanced) and point value
- A1 – 8 challenges, 10% of the final grade, due in 26 days
  - substitution ciphers, hashing, symmetric and asymmetric cryptography, RSA, cryptanalysis
- A2 – 15 challenges, 15% of the grade, due in 24 days
  - network traffic analysis, port knocking, access control, buffer overflow, command injection, format string attack, and SQL injection
- Both A1 and A2 included optional bonus challenges

Scoring – observations

- Some statistically significant correlations between variables mined from CTF games and other forms of the assessment were observed
- Correlation coefficients range only from -0.5 to 0.61
- One of the strongest non-obvious positive correlations:
  the total CTF score of A1 and A2 including bonuses
  and marks from all other types of assessment in the course
  ($r = 0.5$, $p \leq 0.001$)

Scoring – discussion

- Three factors were affecting these results:
- **Total score** of A1 and A2 has a **skewed distribution** (medians were the maximum possible scores without bonus challenges)
- All the **hints** were **for free**, their usage is not reflected in the score
- CTF **portal did not log important game events** (such as displaying the challenge)

Scoring – recommendation for instructors

- Examine what kind of game events are logged by the CTF portal
- Does your CTF portal provide any detailed information about the students' progress (such as the duration of students' interactions with the platform)?

Scaffolding – observations

- Students did not solve some challenges in a reasonable time despite viewing hints
- Only 7 (out of 19) were solved with the median time less than 1 hour after displaying the hint
- After-game survey suggests that some hints did not reach their aim (*Moderately useful* or *Slightly useful*)
- 6 students (out of 20) mentioned hints in their answer to "If you could choose one element of your CTF experience to improve, what would it be?"

Scaffolding – discussion

- Major differences among median times of individual hints indicate some hints were more helpful than others
- Instant feedback provided by students on hints right after solving the challenge in the CTF portal supports this as well
- An example is challenge 2 in A2:
  - the challenge was easy, and hints were clear and guiding,
    e.g., *You need to view the _image file_ to get the flag*
  - all 7 students who provided feedback assessed hints as *Rather Useful*, *Somewhat Useful* or *Useful*

Scaffolding – recommendations for instructors

- Indicate what a hint is about
  - particularly when offering more hints for the same challenge
  - hint cost (if any) is not sufficient
- Pilot test challenge assignments and hints before the game
  - ask TAs or fellow instructors
  - while the challenge assignments can be a bit fuzzy, hints should be clear and straightforward
- Prepare backup hints
  - monitor the ongoing game (submissions, wrong flags, and hint usage) and be ready to add a new hint if needed

Scaffolding – recommendations for developers of CTF frameworks

- Actively offer hints
  - some students tend to beat the challenge without displaying any hints even though the hint may speed up their progress
  - consider adding a feature that will offer a hint to a student after some time of the challenge solving
- Support adaptive hints
  - any step toward dynamically served hints would be beneficial

Plagiarism – observations

- Submissions of the same flag in a time vicinity
  - two challenges solved by 8 and 7 pairs of students within 10 minutes
- Correct flag submitted as an incorrect flag to another challenge
  - 8 cases, time between the submissions varied from 14 seconds to 18 hours
  - two cases where a flag from a still locked challenge was submitted as an incorrect flag to the preceding challenge
- Challenges solved without downloading the file to analysis
  - 11 cases of solving challenges without prior download of a required file
- Quick solves of consecutive challenges
  - 7 very quick solves (9 to 53 seconds vs. minimal possible solve time of 1 min 15 sec)

Plagiarism – discussion

- Time vicinity of submissions may only weakly indicate plagiarism, but quick solves of consecutive locked challenges are more serious evidence
- In A2, instructors questioned such students.
  - **Three of them eventually confessed they used flags shared by their peers.**
  - Some students argued this was only a coincidence since they consulted their approach in a group and then solved the challenge and submitted the flag each on their own.

Plagiarism – recommendations for instructors

- Set rules for students' collaboration during the game in advance
  - Is any discussion about challenges among students forbidden?
  - Communicate these rules clearly and explicitly to students.
- Inform students about how you will check suspicious submissions in advance
  - Instructor may randomly select several students for in-person demonstration.
- Structure related problems to challenge chains
  - It will help not only with revealing plagiarism but also explicitly guide students what must be solved first.

Plagiarism – recommendations for developers of CTF frameworks

- Support challenge chains or dependencies
- Provide built-in analyses for revealing flag sharing
  - All our analyses were done outside the CTF platform using ad-hoc scripts and third-party tools
  - If the results of these analyses were easily accessible in the CTF portal, instructors would benefit from them.

Students' perceptions of the CTF games

- CTF games vs. normal homework assignments
  - **13 students prefer CTF games:** games were fun, hands-on, more interactive, objective, allow to learn to work with security tools, and allow lots of trial and error in exploration
  - 2 normal assignments: games were difficult and very time-consuming
  - 1 is not sure: the normal assignment is little harder to simply copy than CTF games
- For more students' answers regarding hints, immediate feedback, scoreboard, and challenge unlocking see paper (Section 4.4)

Conclusions

- Replacing traditional homework assignments by CTF games is favorable for both instructors and students.
- We highlighted several pitfalls of using CTFs in summative assessment.
- Instructors should carefully consider the game format, scoring, a CTF platform for running the game, and game duration.
- We provide two open-source software plugins for CTFd:
  - https://github.com/janvykopal/ctfd-challenge-feedback
  - https://github.com/janvykopal/ctfd-linear-unlocking
- Full paper is available at https://doi.org/10.1145/3328778.3366893

MUNI C4E

MUNI
C4E

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS

C4E.CZ