Toward a Field Study on the Impact of Hacking Competitions on Secure Development

Daniel Votipka, Hongyi Hu, Bryan Eastes, and Michelle L. Mazurek
SECURE DEVELOPMENT
SECURE DEVELOPMENT
SECURE DEVELOPMENT
SECURE DEVELOPMENT
SECURE DEVELOPMENT

- Non-experts lack experience
- Experts learn through CTFs

[Votipka et al., 2018]
CAPTURE THE FLAG (CTF)

• Attack-oriented competitions
  ▸ Goal: find and exploit vulnerabilities
• Simple, vulnerable programs
• Expose competitors to several classes of vulnerabilities
CAPTURE THE FLAG (CTF)

- Attack-oriented competitions
  - Goal: find and exploit vulnerabilities
- Simple, vulnerable programs
- Expose competitors to several classes of vulnerabilities

Do these help in practice?
RESEARCH QUESTIONS

1. Do CTFs improve prevention of security issues?
2. Do CTFs improve recognition of security issues?
PILOT STUDY OVERVIEW

(1 week)

Dropbox Capture-the-Flag
PILOT STUDY OVERVIEW

Diary Study

Dropbox Capture-the-Flag

Diary Surveys

(1 week)

(10 mins: 6 weeks, 2x/week, 1x/day)
PILOT STUDY OVERVIEW

Diary Study

Dropbox Capture-the-Flag (1 week)

Diary Surveys (10 mins: 6 weeks, 2x/week, 1x/day)

Knowledge Assessment

Pre-CTF Assessment (60 mins)

Post-CTF Assessment (60 mins)
DIARY SURVEYS

• Survey regarding recent commit
  ‣ Issues considered
  ‣ Reasons for considering each issue
  ‣ Actions taken to resolve

• Not security specific

• Open to all Dropbox developers
DIARY SURVEYS

- Survey regarding recent commit
  - Issues considered
  - Reasons for considering each issue
  - Actions taken to resolve
- Not security specific
- Open to all Dropbox developers

Measure impact of CTF on day-to-day decisions
**Knowledge Assessment**

- **Part 1: Find vulnerabilities in insecure code**
  - Copy of the Dropbox codebase
  - 4 known vulnerabilities

- **Part 2: Write a secure program**

- Only CTF participants
KNOWLEDGE ASSESSMENT

• Part 1: Find vulnerabilities in insecure code
  ▸ Copy of the Dropbox codebase
  ▸ 4 known vulnerabilities

• Part 2: Write a secure program

• Only CTF participants

Measure improvements to secure development in a controlled setting
ADDITIONAL METRICS

• Number of flagged commits
• Communication with the Dropbox security team
PILOT PARTICIPATION

• Diary Surveys
  ‣ 28 participants (12 CTF)
  ‣ 169 surveys
• Knowledge Assessment
  ‣ 7 participants
PILOT PARTICIPATION

- Diary Surveys
  - 28 participants (12 CTF)
  - 169 surveys
- Knowledge Assessment
  - 7 participants

- Small sample
- Methodological issues addressed in future iterations
DIARY SURVEYS

• Security considered in 17/124 functionality changes
  ▸ 19% CTF, 13% non-CTF
DIARY SURVEYS

- Security considered in 17/124 functionality changes
  - 19% CTF, 13% non-CTF

CTF participants considered security more often
VULNERABILITIES CONSIDERED

- XSS
- CSRF
- SQLi
- Privacy
- Logic
- Local File Disclosure
- Auth Bug

Percentage of functionality changes

CTF vs Non-CTF
VULNERABILITIES CONSIDERED

Everyone considered logic-based vulnerabilities
VULNERABILITIES CONSIDERED

CTF participants considered non-functionality vulnerabilities from the CTF

Everyone considered logic-based vulnerabilities
REASONS FOR CONSIDERING ISSUES

- Tool
- Teammate
- Standard Practice
- Similar Exp.
- Sensitive Data
- Hacker

Percentage of functionality changes

CTF
Non-CTF

0 10 20 30 40 50 60 70 80
REASONS FOR CONSIDERING ISSUES

CTF participants adopted an adversarial mindset
ACTIONS TAKEN

- Teammate
  - System Doc
  - Previous Exp.
  - Later Review
  - External Doc
  - Expert

Percentage of functionality changes

- Non-CTF
- CTF

Number of Surveys

- 0
- 10
- 20
- 30
- 40
- 50
ACTIONS TAKEN

CTF participants sought help outside of their team.
KNOWLEDGE ASSESSMENT
KNOWLEDGE ASSESSMENT

Average CTF Score

1306
KNOWLEDGE ASSESSMENT

Average Change in Assessment Score: 1.36
KNOWLEDGE ASSESSMENT
Participants with higher than average CTF scores also had higher than average changes in assessment scores.
Participants with higher than average CTF scores also had higher than average changes in assessment scores.
ADDITIONAL METRICS

• Non-CTF participants’ commits were flagged slightly more often
  ▸ 2/17 Non-CTF participants flagged
  ▸ 1/18 CTF participants flagged

• 4 CTF participants alerted security team to potential vulnerability
SUMMARY

1. Do CTFs improve prevention of security issues?
2. Do CTFs improve recognition of security issues?
SUMMARY

1. Do CTFs improve prevention of security issues?
   • Participants who solved more challenges improved in the knowledge assessment
   • Exposure to non-functionality vulnerabilities

2. Do CTFs improve recognition of security issues?
SUMMARY

1. Do CTFs improve **prevention** of security issues?
   - Participants who solved more challenges improved in the knowledge assessment
   - Exposure to non-functionality vulnerabilities

2. Do CTFs improve **recognition** of security issues?
   - Increased consideration of security
   - Improved security team engagement
SUMMARY

1. Do CTFs improve **prevention** of security issues?
   - Participants who solved more challenges improved in the knowledge assessment
   - Exposure to non-functionality vulnerabilities

2. Do CTFs improve **recognition** of security issues?
   - Increased consideration of security
   - Improved security team engagement

Questions:
dvotipka@cs.umd.edu
vulnstudy.cs.umd.edu