Analysis and Exercises for Engaging Beginners in Online CTF Competitions for Security Education

Tanner J. Burns[1], Samuel C. Rios[1], Thomas K. Jordan[1], Qijun Gu[1], Trevor Underwood[2]

[1]Department of Computer Science, Texas State University, San Marcos, TX 78666
Email: {tjb102,scr3,tkj15,qijun}@txstate.edu
[2]Netspend Corporation, Austin, TX 78768
Email: tunderwood@netspend.com

August 15, 2017

Outline

Motivation

- Many students were excited about security and motivated to compete in online CTF competitions.
- BUT,
  - They easily give up after competing in a few online CTF competitions (just like giving up games if they do not see chances of winning and then move on to other games.)
  - They do not have ideas on how to prepare so that they could solve at least one challenge in the competitions.
  - They do not have a set of excises and an easy-to-use platform to practice and develop their security knowledge and skills.
- So, we adopted a data-based approach to design exercises for beginners to grow their interests and skills in computer security.

Contribution

- A study centered around beginner
- Showed the main characteristics of the past security challenges
- Identified the main security issues concerned in the security community
- Highlighted the main knowledge and skills used in the competitions
- Provided a training platform based on PicoCTF
- Deployed a set of exercises based on the analysis of the past security challenges
- Enabled them to study and practice by themselves

Outline

Platform Choices

- Platform
  - Open source
  - Standalone package
  - Easy deployment
- Functionality (for beginners):
  - User management, web interface, problems setup, problems grading, statistics of players and teams, and so on.
- Development (for developers and administrators):
  - Deployment: Vagrant, Docker, Native
  - Coding Language: Python, PHP
  - Documentation: Installation, Adding exercises and features

Platform Choices

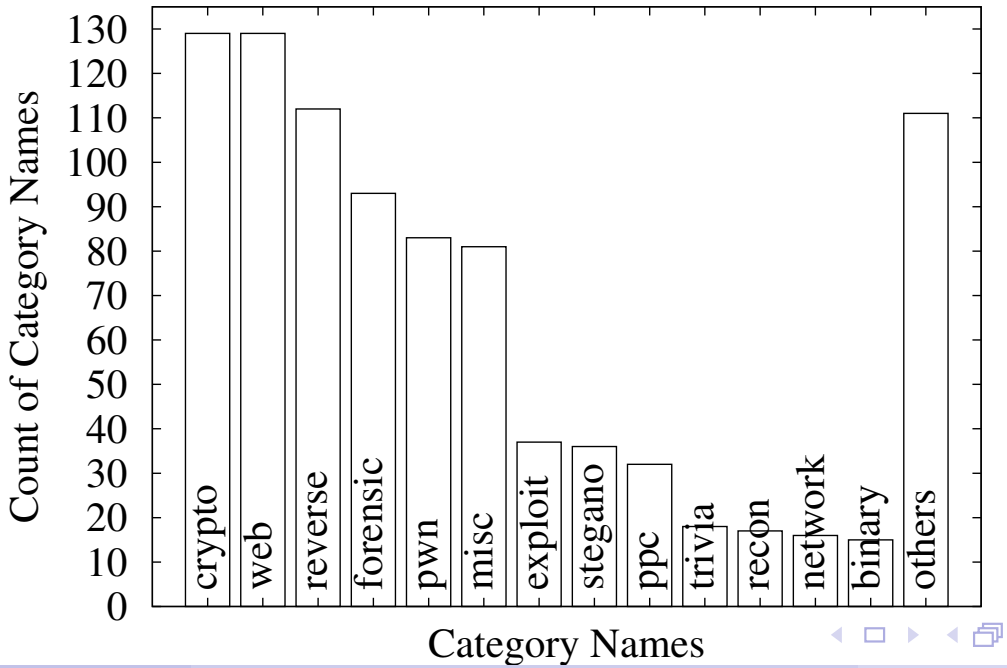
- Sources: hosted on github from past competitions
- Our choice: PicoCTF

Table: Comparison of CTF Platforms

	Installation	Language	Documentation
PicoCTF	Vagrant ✓	Python ✓	Good ✓
OpenCTF	Docker ✓	Python ✓	Simple
CTFd	Native	Python ✓	Simple
FbCTF	Vagrant ✓	PHP	Good ✓
TinyCTF	Native	Python ✓	Simple

Exercise Categories

- First group: chosen for our exercises
  - Crypto, Web, Reverse, Forensic, Pwn, and Misc
- Second group
  - Exploit, Stegano, Ppc, Trivia, Recon, Network, and Binary
- Others: sixty-four category names
  - For example: Admin, Unknown, Mobile, Coding, Joy, Shellcode, Cgc, Crack, Grab bag, ...



Difficulty Levels

- Three levels: easy, medium, and hard.
- Three quarters of the security challenges are at the easy and medium levels.
- Our exercises are based on the easy and medium security challenges for beginners.

Category	Easy	Medium	Hard
crypto	192 (48%)	129 (32%)	83 (21%)
web	152 (41%)	150 (40%)	70 (19%)
forensic	263 (50%)	186 (35%)	79 (15%)
reverse	77 (22%)	131 (38%)	136 (40%)
pwn	66 (19%)	138 (39%)	148 (42%)
misc	96 (48%)	67 (34%)	35 (18%)
total	846 (38%)	801 (36%)	551 (25%)

Outline

Summary of Exercises

- PicoCTF-derived standalone virtual box image
- Six categories of exercises with thirty-five exercises
- All exercises adopted from past CTF competitions
- Half at the easy level and half at the medium level
- Partial and full solutions

Coding Exercises (Misc)

- To improve programming proficiency
  - Automated data processing
  - Programmatically data analysis
  - Utilizing libraries and tools (such as Python's)
- Coding exercises: must-have coding skills
  - Number and string conversion: hexadecimal and binary conversions, string and number conversions, large number arithmetic, Base64 encoding and decoding, string splitting and concatenation
  - File manipulation: open, read, process, write
  - Networking: create services, make and send arbitrary packets to remote servers, and process packets received from remote servers

Cryptographic Exercises

- Ciphers: symmetric (Caesar, Vigenere and AES), asymmetric (RSA), and hash (MD5 and SHA1)
- Programming: encryption, decryption, cryptographic analysis, substitution, factorization, hash collision, and so on

Groups	Problem Ratios	Cipher Counts	Top 2 Ciphers
Custom	37.3%		XOR
Symmetric	34.4%	36	AES, Caesar
Asymmetric	21.3%	10	RSA, ECC
Hash	5.3%	5	MD5, SHA1/2
Misc	1.7%	4	DSA, SSL

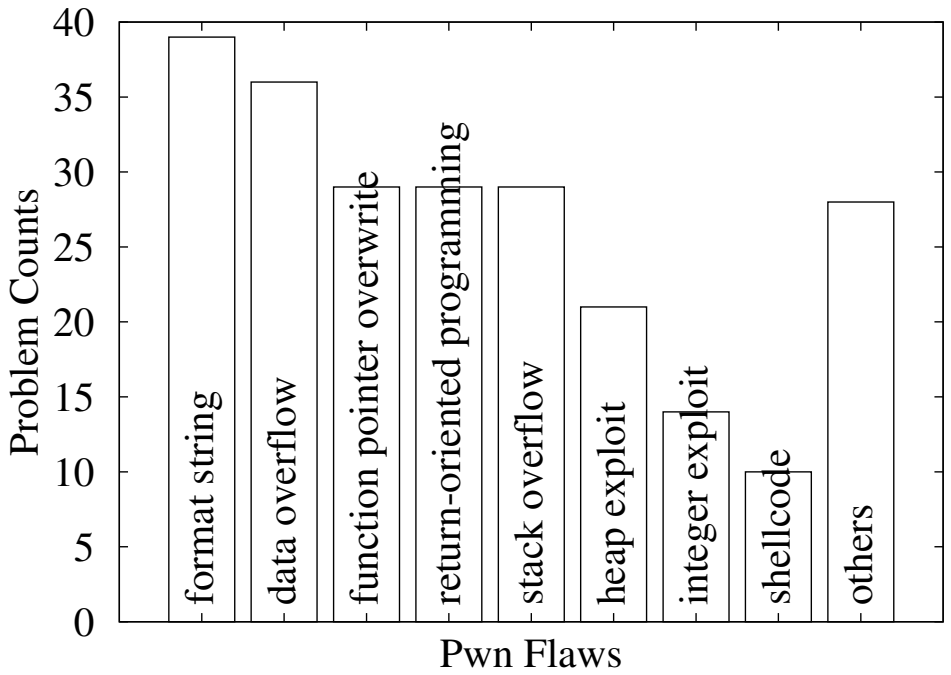
Reverse Engineering Exercises

- Static analysis: disassemble and decompile
  - X64 and X86 binaries
  - Java and Android applications
- Dynamic analysis
  - Tracing library calls and system calls, and overloading library functions
  - Debugging techniques: (1) stepping and breaking execution, (2) watching and changing variables, memory and registers

Rank	Reverse		Pwn	
1	X64	33.9%	X64	31.7%
2	X86	28.7%	X86	31.3%
3	Java	9.1%	C	16.7%
4	PE32	6.3%	Python	8.2%
5	Python	4.5%	Bash	3.2%
6	Others	17.5%	Others	8.9%

Pwn Exercises

- Format string
- Overflow: data overflow, stack overflow, heap overflow and integer overflow
- Function pointer overwriting
- Return-oriented programming (ROP)



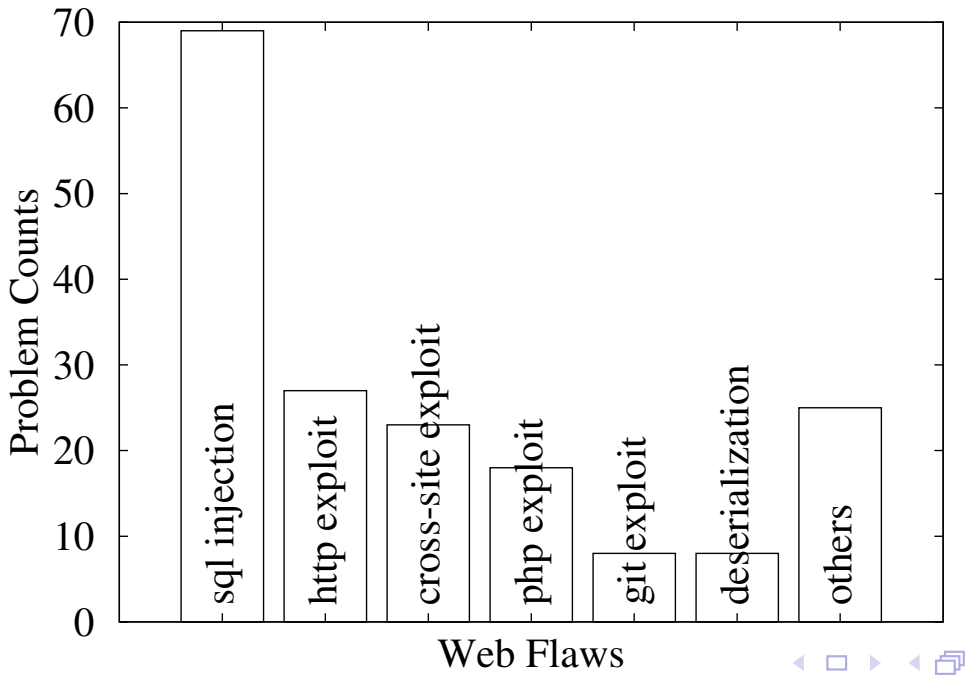
Forensic Exercises

- Image: png and jpg
- Network trace: pcap
- Multimedia: wav
- Data file: zip, text, pdf

Group	Format Counts	Top 2 Formats
image	8	png, jpg
network	5	pcap, tcpdump
audio	5	wav, mp3
disk*	6	dd, ext4
archive	7	zip, tar
dump*	6	memory, vbox
text	6	text, c, html
pdf	1	pdf

Web Exercises

- To proficiently use CURL and web development tools
- To inspect web pages and web traffics
- To inspect and manipulate cookies, sessions, URLs, form data, JSON data and web agents on the client side
- To exploit the top three flaws: SQL injection, http exploit, and cross-site script exploit.



Outline

Class Settings

- 46 beginner students
- CTF exercises were a part of individual homework assignments
- Provided partial solutions on the key steps
  - Guide the students with example techniques on the most challenging steps
  - Ask the students to figure out the missing steps and complete the exercises by themselves
- Minimum intervention from the instructor
  - To enable beginners to learn and practice by themselves
  - To enable beginners to build technical and psychological confidence by themselves
- Anonymous survey to get feedback from the students
- Goals: to identify issues and assess appropriateness of the exercises for beginners

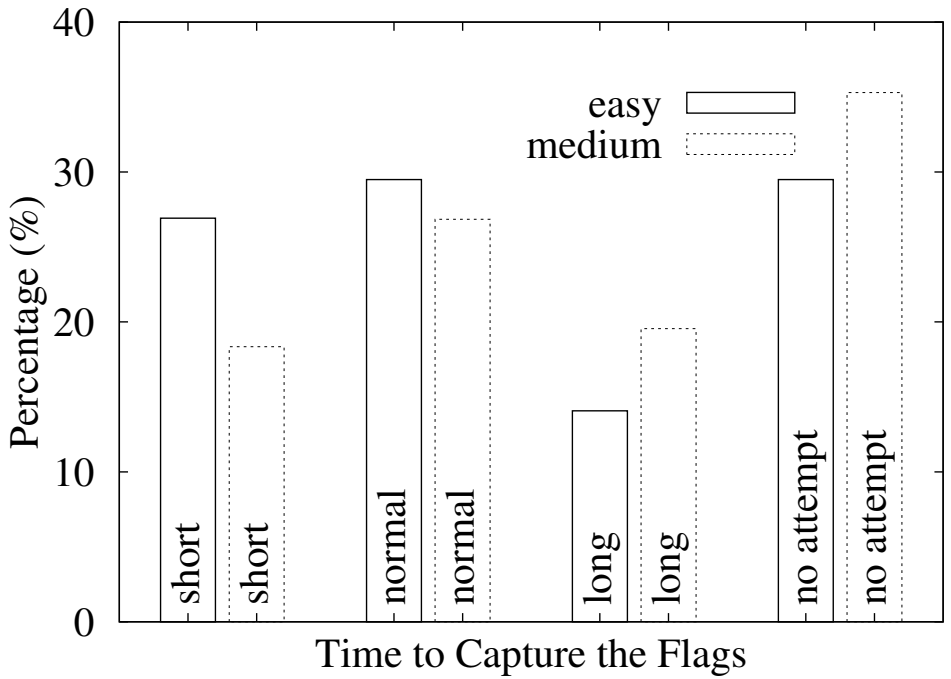
Observations

Usage issues:

- Several students were using tablet computers that do not support Virtual Box.
- Many students did not have enough computer administration skills to install and setup the needed tools and libraries.
  - Many tools and libraries are Linux-based.
  - Many students use Windows and Mac OS X.
- About 13% of students gave up on the exercises due to these issues.

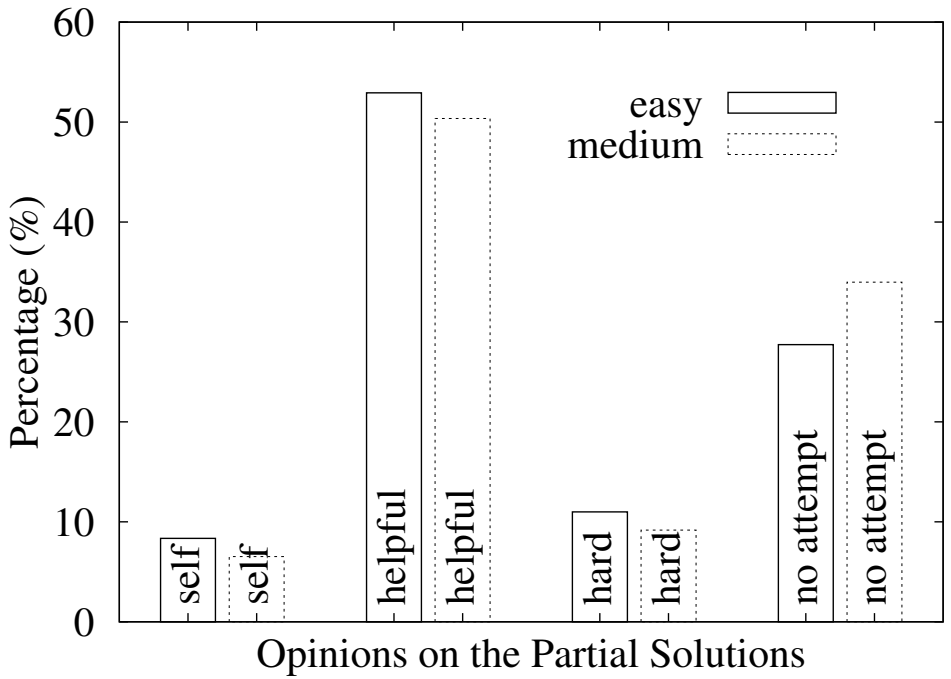
Observations

- Time to complete the exercises (assessing the difficulty levels)
  - Short: work time fewer than 15 minutes
  - Medium: work time between 15 and 40 minutes
  - Long: work time greater than 40 minutes
  - No attempt: did not work on the exercises



Observations

- Provided partial solutions of the exercises
  - Self: got the flags without reading the solutions
  - Helpful: got the flags with the help of the solutions
  - Hard: did not understand the solutions
  - No attempt: did not work on the exercises



Outline

Conclusion and Future Works

- A downloadable standalone CTF package for beginners to use by themselves
- A set of exercises in six categories with must-have skills for beginners
- Positive feedbacks from beginner students
- Future works
  - Missing skills of system administration and management
  - Missing defensive techniques and skills
  - Performance of beginners in CTF competitions after studying these exercises