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

Tanner J. Burns¹, Samuel C. Rios¹, Thomas K. Jordan¹, Qijun Gu¹, Trevor Underwood²

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

August 15, 2017
Outline

1. Motivation and Contribution
2. Setup
3. Exercises
4. Lessons Learned
5. Conclusion and Future Works
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 exercises 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

1. Motivation and Contribution
2. Setup
3. Exercises
4. Lessons Learned
5. Conclusion and Future Works
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

<table>
<thead>
<tr>
<th>Platform</th>
<th>Installation</th>
<th>Language</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PicoCTF</td>
<td>Vagrant ✓</td>
<td>Python ✓</td>
<td>Good ✓</td>
</tr>
<tr>
<td>OpenCTF</td>
<td>Docker ✓</td>
<td>Python ✓</td>
<td>Simple</td>
</tr>
<tr>
<td>CTFd</td>
<td>Native</td>
<td>Python ✓</td>
<td>Simple</td>
</tr>
<tr>
<td>FbCTF</td>
<td>Vagrant ✓</td>
<td>PHP</td>
<td>Good ✓</td>
</tr>
<tr>
<td>TinyCTF</td>
<td>Native</td>
<td>Python ✓</td>
<td>Simple</td>
</tr>
</tbody>
</table>
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, ...

Count of Category Names

Category Names

- crypto
- web
- reverse
- forensic
- pwn
- misc
- exploit
- stegano
- ppc
- trivia
- recon
- network
- binary
- others
## 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.

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

1. Motivation and Contribution
2. Setup
3. Exercises
4. Lessons Learned
5. Conclusion and Future Works
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
Exercises

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

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

<table>
<thead>
<tr>
<th>Rank</th>
<th>Reverse</th>
<th>Pwn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X64</td>
<td>33.9%</td>
</tr>
<tr>
<td>2</td>
<td>X86</td>
<td>28.7%</td>
</tr>
<tr>
<td>3</td>
<td>Java</td>
<td>9.1%</td>
</tr>
<tr>
<td>4</td>
<td>PE32</td>
<td>6.3%</td>
</tr>
<tr>
<td>5</td>
<td>Python</td>
<td>4.5%</td>
</tr>
<tr>
<td>6</td>
<td>Others</td>
<td>17.5%</td>
</tr>
</tbody>
</table>
Pwn Exercises

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

Problem Counts

<table>
<thead>
<tr>
<th>Pwn Flaws</th>
<th>Problem Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>format string</td>
<td>40</td>
</tr>
<tr>
<td>data overflow</td>
<td>35</td>
</tr>
<tr>
<td>function pointer overwrite</td>
<td>30</td>
</tr>
<tr>
<td>return-oriented programming</td>
<td>25</td>
</tr>
<tr>
<td>stack overflow</td>
<td>20</td>
</tr>
<tr>
<td>heap exploit</td>
<td>15</td>
</tr>
<tr>
<td>integer exploit</td>
<td>10</td>
</tr>
<tr>
<td>shellcode</td>
<td>5</td>
</tr>
<tr>
<td>others</td>
<td>0</td>
</tr>
</tbody>
</table>

Tanner J. Burns, Samuel C. Rios, Thomas IAnalysis and Exercises for Engaging Beginners in Online CTF Competitions for Security Education
Forensic Exercises

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

<table>
<thead>
<tr>
<th>Group</th>
<th>Format Counts</th>
<th>Top 2 Formats</th>
</tr>
</thead>
<tbody>
<tr>
<td>image</td>
<td>8</td>
<td>png, jpg</td>
</tr>
<tr>
<td>network</td>
<td>5</td>
<td>pcap, tcpdump</td>
</tr>
<tr>
<td>audio</td>
<td>5</td>
<td>wav, mp3</td>
</tr>
<tr>
<td>disk*</td>
<td>6</td>
<td>dd, ext4</td>
</tr>
<tr>
<td>archive</td>
<td>7</td>
<td>zip, tar</td>
</tr>
<tr>
<td>dump*</td>
<td>6</td>
<td>memory, vbox</td>
</tr>
<tr>
<td>text</td>
<td>6</td>
<td>text, c, html</td>
</tr>
<tr>
<td>pdf</td>
<td>1</td>
<td>pdf</td>
</tr>
</tbody>
</table>
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

1. Motivation and Contribution
2. Setup
3. Exercises
4. Lessons Learned
5. Conclusion and Future Works
Lessons Learned

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
Lessons Learned

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

1. Motivation and Contribution
2. Setup
3. Exercises
4. Lessons Learned
5. Conclusion and Future Works
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