Automatic Problem Generation for Capture-the-Flag Competitions

Jonathan Burket  
jburket@cmu.edu

Peter Chapman  
peter@cmu.edu

Tim Becker  
tjbecker@cmu.edu

Chris Ganas  
cganas@cmu.edu

David Brumley  
dbrumley@cmu.edu

Carnegie Mellon University

This material is based upon work supported by the National Science Foundation under Grant No. 1419362 and by a Graduate Research Fellowship under Grant No. 0946825.

8/11/2015
Outline

Overview of Automatic Problem Generation
Flag Sharing in picoCTF 2014
Future Work
“Jeopardy-Style” Capture-the-Flag

Cryptography
100

Cryptography
200

Web
100

Web
200

Forensics
100

Flag:
“sql_injection_rox”

Points!
Example CTF Problem

You find an encrypted message written on the documents. Can you decrypt it?
“rfc qcapcr nyqqnfpyqc gq bmuasiugssaxxlextkasoklnctcidhm”

“the secret passphrase is dowcukwiuuczzngzvmcuqmpvekfjo”
Flag Sharing

Flag: "sql_injection_rox"

Alice → Flag 
Bob → Flag

Points!

Points!
Flag Sharing

Flag: “sql_injection_rox”

Alice → Flag Sharing

Bob, Carol, Dave → Points!
Changing Problems Automatically

“the secret passphrase is dowcukwiuucczzngzvmcuqmnvpvekfjo”

k=24  “rfc qcapcr nyqqnfpypyc gq bmuasiugssaxxlextkasoklntcidhm”

“the secret passphrase is vrsoblzffauncgrgknleuxedsknnhb”

k=24  “rfc qcapcr nyqqnfpypyc gq tpqmzjxddylaepeiljcsvcbqillfz”
Changing Problems Automatically

“the secret passphrase is dowcukwiuuczzngzvmcuqmnpvekfjo”

$\mathbf{k=24}$ $\Rightarrow$ “rfc qcapcr nyqqnfpyqc gq bmuasiugssaxxlextkasoklnctcidhm”

“the secret passphrase is dowcukwiuuczzngzvmcuqmnpvekfjo”

$\mathbf{k=6}$ $\Rightarrow$ “znk ykixkz vgyyvnxgyk oy juciaqcoaaiffitmfbisiawstvbkqlpu”

These different versions can be generated automatically!
Automatic Problem Generation

Through *Automatic Problem Generation*, different competitors can receive different versions ("*autogen* problem instances") of a given problem.

Alice

“wkh vhfuhw sdrvskudvh lv yuvreociidxqfjujnqohxahgynqke”

Flag: “vrsoblzffauncggrkknleux…”

Bob

“rfc qcapcr nyqqnfpyqc q gq bmuasiugssaxxlkxkasoklntcidhm”

Flag: “dowcukwiiuuczngzvmc…”

Flags can’t be shared!
Templated Autogen Problems

“What is the value of {register} after executing the instruction at address {memory address}?"

“What is the value of eax after executing the instruction at address 0x12345678?”

“What is the value of ebx after executing the instruction at address 0x20202020?”

Problem instances are essentially the same, but have slightly different details.

Good for detection and prevention of flag sharing
We detect **flag sharing** by looking for cases where users submit incorrect flags that are correct for an instance other than their own.
distinct, per instance flags allows competition to detect the source of shared flags
Challenges

• Balanced Difficulty
• Bug Prevention
• Scalability and Deployment
Balanced Problem Difficulty

“the secret passphrase is dowcukwiuucczzngzvmcuhmqnpevkefjo”

$k=0$ “the secret passphrase is dowcukwiuucczzngzvmcuhmqnpevkefjo”

Need to make sure each problem instance is reasonably close to the same difficulty.
Scalability and Deployment

• Each team must be given a team-specific problem instance
• Problem instance generation must not bottleneck the performance of the competition
• For problems served from remote machines, instances must be deployed to those machines and synchronized with the grading server
Automatic Problem Generation in picoCTF 2014
picoCTF 2014

- 3,185 Eligible Teams
- 9,738 Eligible Students
- 12 Day Competition
- 66 Security Challenges
- Cash Prizes for Top 10 Teams
Automatically Generated Problems in picoCTF 2014

<table>
<thead>
<tr>
<th>Problem Name</th>
<th>Category</th>
<th>Score</th>
<th>Teams Solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tyrannosaurus Hex</td>
<td>Miscellaneous</td>
<td>10</td>
<td>3185</td>
</tr>
<tr>
<td>No Comment</td>
<td>Web</td>
<td>20</td>
<td>2952</td>
</tr>
<tr>
<td>Caesar</td>
<td>Cryptography</td>
<td>20</td>
<td>2648</td>
</tr>
<tr>
<td>Internet Inspection</td>
<td>Web</td>
<td>30</td>
<td>2704</td>
</tr>
<tr>
<td>Javascrypt</td>
<td>Web</td>
<td>40</td>
<td>1719</td>
</tr>
<tr>
<td>Substitution</td>
<td>Cryptography</td>
<td>50</td>
<td>1677</td>
</tr>
<tr>
<td>ZOR</td>
<td>Cryptography</td>
<td>50</td>
<td>554</td>
</tr>
<tr>
<td>Basic ASM</td>
<td>Reversing</td>
<td>60</td>
<td>887</td>
</tr>
<tr>
<td>Repeated XOR</td>
<td>Cryptography</td>
<td>70</td>
<td>182</td>
</tr>
<tr>
<td>Block</td>
<td>Cryptography</td>
<td>130</td>
<td>35</td>
</tr>
</tbody>
</table>
Flag Sharing in picoCTF 2014

127,412 flags (correct and incorrect) submitted to automatically generated problems
1,081 (0.84%) labeled as “shared flags”
530 distinct flag sharing cases
14% (460/3185) of teams involved
  87% of these teams shared flags for only one problem
68% teams attempting to submit a shared flag eventually solved the problem correctly
Likely Source of Shared Flag
Likely Source of Shared Flag

Unknown, 126, 25%
Likely Source of Shared Flag

- **Unknown, 126, 25%**
- **Other Online Disclosure, 12, 2%**
- **Stack Overflow, 47, 10%**
- **Youtube, 47, 10%**
Hey can you the javascript - 40, basic asm 60, common vulnerability exercise -20, ? Unless anyone already know the answer? if u do please comment asap. and please can you do it!

Hide replies

4 hours ago
picoclf 2014 challenge javascript 40
Reply · I like · P I

4 hours ago
Commom 20 so easy, so i don't make video
Reply · I like · P I

4 hours ago
Flag_2238 doesn't work for me.
Reply · I like · P I

3 hours ago
U sould do as that video
Reply · I like · P I
Likely Source of Shared Flag

- Unknown, 126, 25%
- Other Online Disclosure, 12, 2%
- Stack Overflow, 47, 10%
- Youtube, 47, 10%
Likely Source of Shared Flag

- Unknown, 126, 25%
- Other Online Disclosure, 12, 2%
- Submitted by Another Team at School (Earlier), 264, 53%
- Youtube, 47, 10%
- Stack Overflow, 47, 10%
Time Between Original Submission and Copied Flag Submission

Percent of Submissions (Cumulative)

Number of Hours after Original Submission
Future of Automatic Problem Generation in CTFs
Automatically Generating Problems and Solutions for Natural Deduction

Unair Z. Ahmed
ITC Karjat
unair14123@gmail.com

Sumit Gulwani
MSR Redmond
sumit@microsoft.com

Ameey Karlaare
ITC Karjat
karlaare@cece.ernet.in

ABSTRACT

Natural deduction, although a method for establishing validity of propositional type argument, helps develop important skills. Such skills are taught in a limited and indirect manner, namely through practice exercises or problem generation, which are hard to generate and pose problems in a way that are relevant, difficult and interesting to the user. We present a technique that is designed to efficiently generate difficult and interesting problems for natural deduction.

A B N

Automating Exercise Generation: A Step towards Meeting the MOOC Challenge for Embedded Systems

Dong Cai
dgc@CMU.edu

Ben Sheats
bks9@cmu.edu

Rohit Singh
MIT-CRL

Venkatesh Ganesan
ganesanch@cmu.edu

Sumit Gulwani
Microsoft Research

Satnam Rajan
Microsoft Research

ABSTRACT

The current state of automated exercise generation in computer science education is very limited. There is a need for an automated exercise generation tool to help instructors and students in the process of acquiring knowledge and skills in computer science. We present an approach to automatically generate exercises for MOOC platforms.

A B N

Automatically Generating Algebra Pre-Calculating

Sheng J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.

Seth J. A.
void check_input(int v0, int v1) {
    v0 += 7;
    int v2 = 0;
    v0 *= v2;
    v1 += 2;
    if(v0 < 45) {
        int v3 = v0;
        v2 -= v3;
    }
    v1 *= 8;
    int v4 = 3;
    v4 *= v1;
    if(v0 != 0) {
        you_lose();
        return;
    }
    if(v2 != 0) {
        you_lose();
        return;
    }
    if(v4 != -456) {
        you_lose();
        return;
    }
    get_key();
    return;
}

void check_input(int v0, int v1, int v2, int v3, int v4, int v5, int v6) {
    int v7 = -1;
    v3 -= 3;
    v6 -= 1;
    for(int v8=1; v8<= 9; v8+=v1) {
        v5 += v8;
        v2 *= v6;
    }
    int v9 = 6;
    int v10 = 0;
    if(v2 <= v9) {
        v5 *= v4;
        v0 += v3;
    } else {
        v3 |= v6;
    }
    int v11 = v6;
    v11 += v10;
    for(int v12=1; v12< 10; v12+=v5) {
        int v13 = v7;
        v6 *= v0;
        if(v5 == 44) {
            v13 -= v2;
            v11 *= v10;
        } else {
            v13 *= v11;
        }
        get_key();
        return;
    }
}
Synthesized Autogen Problems

Easy

Hard
Questions?

https://github.com/picoCTF/picoCTF-Platform-2