AN OFFLINE CAPTURE THE FLAG-STYLE VIRTUAL MACHINE FOR CYBER SECURITY EDUCATION

Tom Chothia
Chris Novakovic

University of Birmingham

Introduction

- A VM to support cyber security education.
- The VM creates unique flags, which the students get if they solve the exercise.
- Flags submitted by students online and used to support marking.
- Comparison of flag only marking vs. traditional continuous assessment marking.

Capture the Flag Competitions



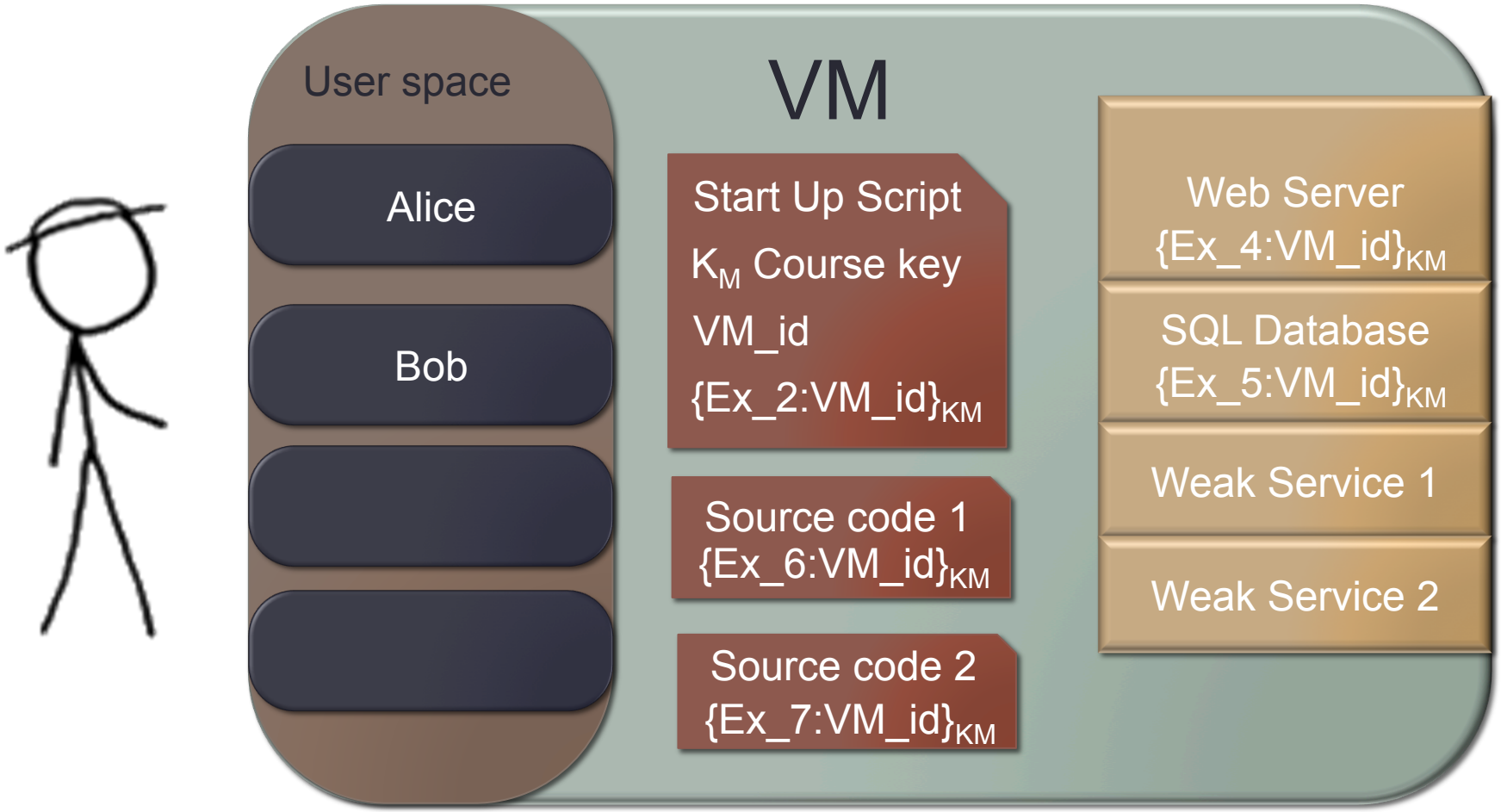
Flag We Have Known:

- Flag:'a012c434d1ec6db911fda4884de14fdd’
- 31c3_a5bb3ead8fbc6617374ea3f57f0563d2
- the_snozberries_taste_like_snozberries
- 'this is a key’
- SECCON{TeaBreakAtWork}
- FLAG(CYBERCONTROLCYBERDATACYBERCORPORATION)

Problems with using CTFs in Education

- Everyone gets the same flag and same infrastructure.
- Short lived intense competitions
  - We don’t want students to “win” or “loose” the course.
  - Not compatible with an 11 week module.
- Live, online competitions use complex infrastructure and need a lot of support.
- IT Services do ***not*** support large amounts of hostile attack traffic on their network.

Our Solution



Exercises

- Encryption
  - Students need to write code to decrypt RSA & AES: ECB, CBC, CTR, CCM mode encryption.
  - Flags are inside the encrypted messages.
  - Perform known plaintext bit flipping attacks.

- Access control
  - Start up script writes flags to files and sets access permissions.
  - Subtle errors in the access control allow access to these flags and the shadow file
    - E.g. chained confused deputy using setuid.
  - Cracking the passwords gives access to account with flag
  - Root & some other passwords cannot be cracked.

Examples: Protocol Questions

1. $C \rightarrow S$: "Connect Protocol 1"
2. $S \rightarrow C$: $\{N_s\}_{Kcs}$
3. $C \rightarrow S$: $\{N_c\}_{Kcs}$
4. $S \rightarrow C$: $\{N_c, N_s\}_{(Ns \oplus Nc)}$
5. $C \rightarrow S$: $\{N_s, N_c\}_{(Ns \oplus Nc)}$
6. $S \rightarrow C$: $\{\text{Secret Value}\}_{(Ns \oplus Nc)}$

Student Feedback

- Each time taught it has been the most popular course in the school

The module content is very interesting and the exercises provide good application of the content.

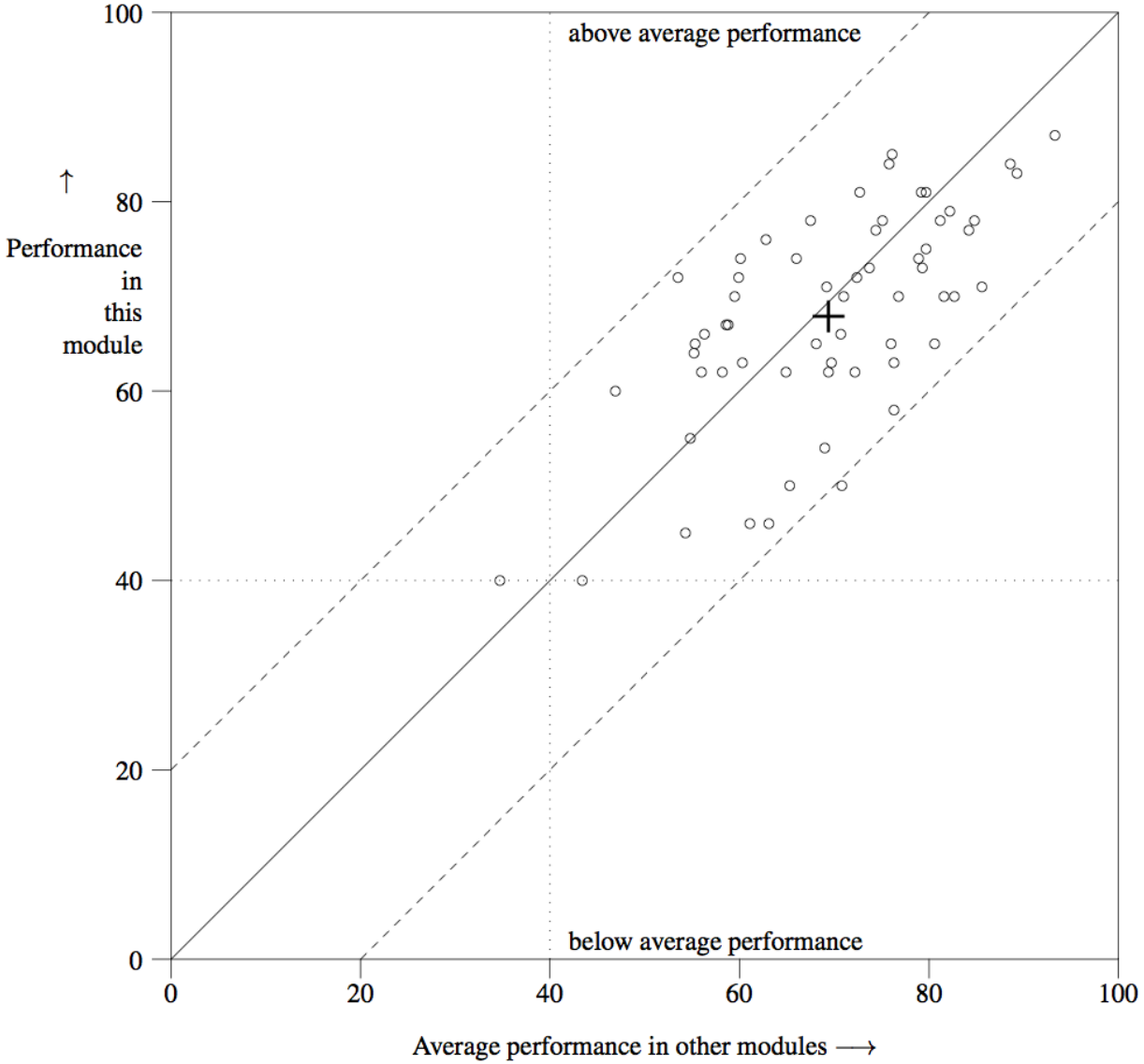
By far the most interesting, fun and enjoyable module I've been enroled on at University, cannot fault it.

CONTENT IS INCREDIBLY WORTHWHILE

Exercises are fun and challenging

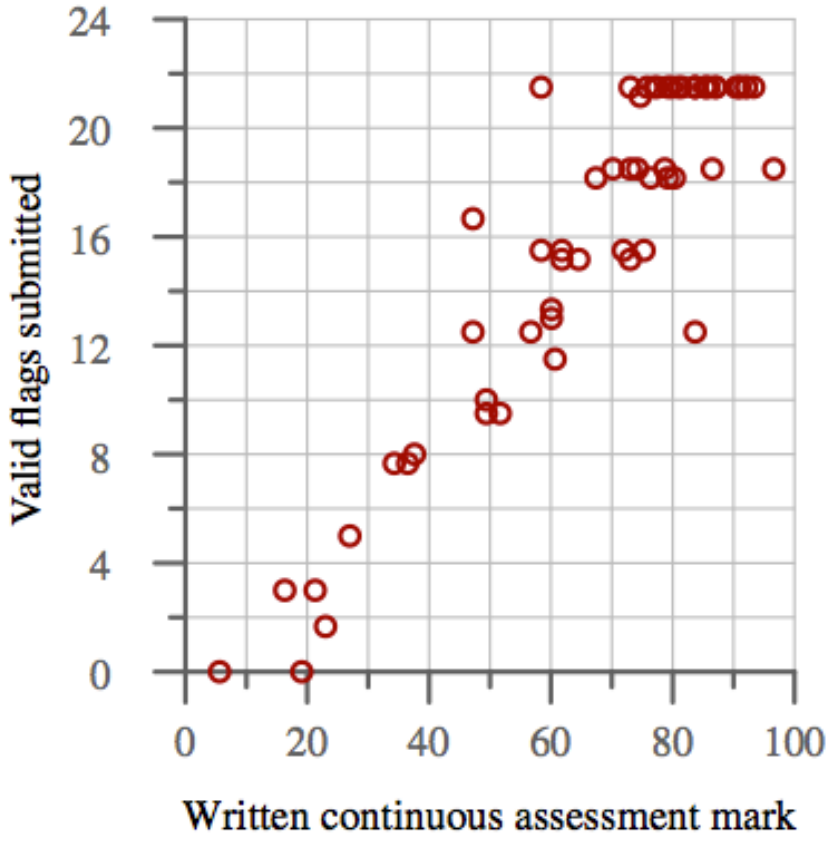
Challenging yet interesting content. Brilliant module overall!

It was all pretty damn cool.

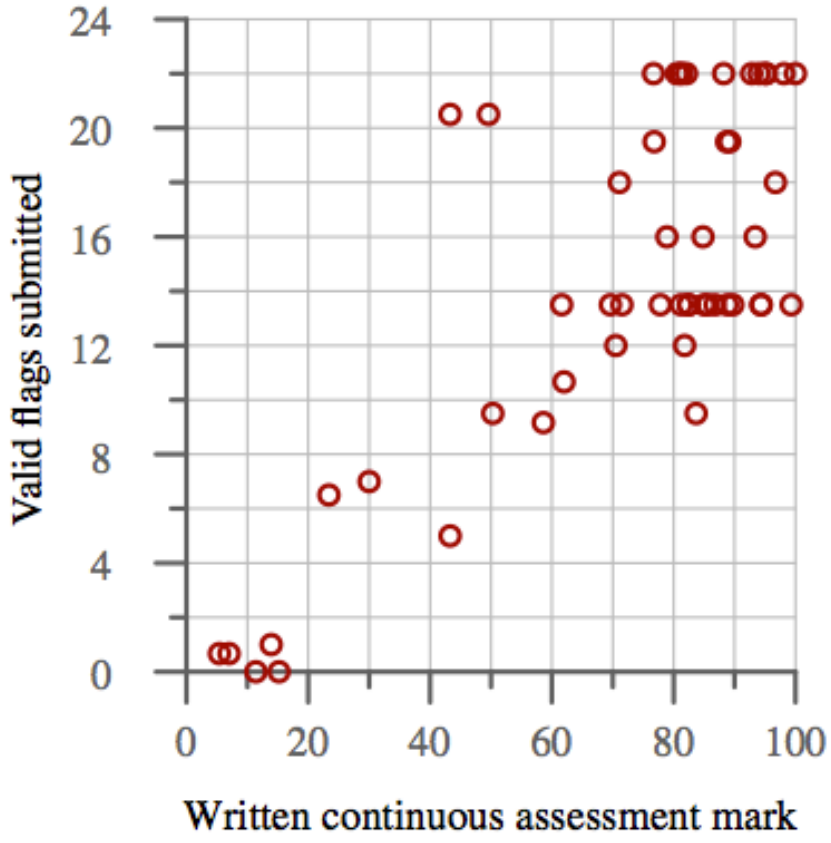

100
above average performance
↑
80
Performance
in
this
module
60
40
20
0
below average performance
0
20
40
60
80
100
Average performance in other modules ⟶

Assessment

- We used the existence of flags to aid human marking of written answers and code.
- Flags were required and gave the marker confidence that code worked and that written answers were correct.
  - But we gave no marks for just the flags
- Would fully automated marking based only on flags have worked?
  - Full marking of exercises let's us test this.



(a) Spring 2014



(b) Autumn 2014

	Spring 2014	Autumn 2014
All marks	0.92	0.84
Marks ≤ 70	0.37	0.22
Marks < 70	0.91	0.82

Flag marking

- Difference between a 1st and a very high 1st was showing a full and complete understanding. E.g.
  - Protocols, low 1st: correct, but inefficient attacks and fixes
  - Protocols, high 1st: most efficient attacks and fixes possible
  - Reverse engineering, low 1st: right answer found by lots of trial and error.
  - Reverse engineering, high 1st: complete understanding of the assembly.
- Fully automated flag only marking would be fine for broad classifications, but would not recognize the best students.
- Better than no assessed course work.

Using Flags to Assess Ability

- Could “better” flag exercises accurately assess ability?
  - As exercises become harder, luck plays more of a part.
  - Too many exercises make them a race against time.

- IMHO: Flags are a great way to make CTFs fun, but can only play a small part in academic assessment.

- Side note: the best CTF teams will write up their CTF results.

Plagiarism

- Plagiarism is an issue with all coursework.
- 4 cases of plagiarism on 3 iterations of the course.
  - Four submitted the same flags & almost the same answers.
  - Two students submitted different flags & almost the same answers.
  - Two students submitted the same flags.
  - Another two students submitted the same flags.
- Roughly half the rate of plagiarism as before flags.
- Questioned informally, students said they didn't plagiarise flags because they had to do written answers and vice versa.

Other issues

- Other problems with flag only marking:
  - Value of written solutions and feedback
  - Written work stops flag plagiarism
- Online monitoring of token submissions.
  - Very useful.
- Students can bypass system
  - Good students that finish exercises early, encouraged to do this
  - Designed to be harder to bypass exercises than solve correctly.

Summary

- Design for offline CTF-style, cyber security education VM.
  - Each student's VM contains unique flags.
  - Very popular with students but still difficult exercises.
  - Publically available, no restrictions on use
  - http://www.cs.bham.ac.uk/~tpc/SecEduVM/
- Flags support human marking and discourage plagiarism.
- Marking based only on flags would not have accurately assessed good students.