A "Divergent"-themed CTF and Urban Race for Introducing Security and Cryptography

Wu-chang Feng
Portland State University
Department of Computer Science

Abstract

There is a recognized shortage of students who are interested in learning computer and network security. One of the underlying reasons for this is a lack of awareness and motivation to study the subject. In order to tackle this problem, we have developed an introductory cryptography and security curriculum that attempts to inspire students to pursue this career path.

Towards this end, the curriculum we have designed motivates the importance of the field and contains a variety of activities intended not only to teach students basic concepts, but also allow them to develop technical skills in a fun and engaging manner. In particular, we employ a novel set of capture-the-flag (CTF) exercises and a physical activity based on an urban race, both of which are tied into a fictional story that students act out. The storyline follows a book series that many young adults of this generation are familiar with: the Divergent books written by Veronica Roth [1]. Using this approach, we have successfully delivered our curriculum at multiple schools throughout Oregon.

1 Introduction

With concerns over securing critical cyber-infrastructure growing, there is an emerging need for highly-skilled security practitioners. As with many skills in life such as athletics and musical instruments, it is often the case that elite levels of performance are reached when one starts early and sustains practice over a long period of time [2]. There has been a significant push towards introducing programming and computer science into grade schools with summer camps and events such as the Hour of Code [3] and Lego Robotics. While enrollment in Computer Science has swelled, interest in the computer security sub-specialty has not.

There have been initial efforts that address this problem. Courses and camps such as NICERC's CyberDiscovery camps, NSA/NSF's GenCyber camps, and Saturday Academy's CyberAcademy camp seek to provide compelling content and experiences that engage high-school students early. In addition, "Capture-the-Flag" (CTF) security competitions such as picoCTF and hsCTF seek to provide introductory venues for students to build their security skills. Both formats seek to inspire students to discover more about the topic and work to intrinsically motivate them to continue.

Towards this end, this paper describes an integrated curriculum, a set of CTF challenges, and an Urban Race for use in cybersecurity-themed camps to introduce security to students with no prior exposure to security, programming, or cryptography. The learning vehicle combines puzzle solving, a storyline that is based on a popular young adult book series, and a physical game, to intrinsically motivate students to solve security challenges. By creating a positive and memorable experience, we aim to make the activities something all students are excited to undertake. The vehicle uses CTFs to tap into the intrinsic reward that people get in solving puzzles, much in the same way Sudoku and crossword puzzles do. It uses a popular storyline in order to tap into the emotional connection students have with familiar characters so that they are motivated to find out how the story ends. Finally, to wrap up the camp, it uses a physical game that drops participants into the storyline itself, allowing them to become full participants in the story and giving them control over how the story ends. Throughout each activity, security concepts are embedded and used with specific purposes and in memorable contexts in order to help students retain information.

2 Design

The main goal of the curriculum is to introduce data encoding and cryptography to students with no prior experience via a combination of short lectures and numerous

hands-on activities. In addition, we aim to introduce key concepts in attacking and defending computer systems so that students can develop an appreciation of the importance of the area and some familiarity with common security-related concepts and terms. To meet these objectives, we use a number of novel mechanisms.

2.1 Cryptography curriculum

The main technical content and skills we seek to develop in students focus on data encoding and cryptography. The curriculum is split across five, 1-hour modules and assumes that students have no prior experience with programming or computer security.

The first module of our curriculum sets the motivation for learning the technical material by describing the importance cryptography throughout history and how it has changed the course of it. Historical stories used include the execution of Mary, Queen of Scots, the Zimmerman telegram, "Stalin's spies", and the Enigma machine. Mixed in with historical examples are relevant current events with global impact such as Stuxnet's use of compromised certificates as well as local impact such as the use of Firesheep to hijack social media accounts [4].

Because cryptography is now mostly done in the digital domain, the second module begins by teaching students the binary and hexadecimal number system for representing digital information. From this, students then examine encodings that map numeric values into non-numeric information such as the encoding of the alphabetic characters in ASCII. To show how information can be represented in arbitrary ways, the underlying workings of visual encoders such as 1-D barcodes and 2-D QR codes are examined. The encoding unit finishes by examining ways in which information can be hidden via steganography with examples that include placing information in image metadata to transmit messages covertly.

The third module of our curriculum picks up on the idea of sending covert messages by introducing simple transposition ciphers and monoalphabetic substitution ciphers. Ciphers covered include Caesar, Scytale, columnar transposition, and simple substituion. For each cipher, the computational difficulty of breaking them using brute-force methods or character frequency analysis is examined. This leads the students to the polyalphabetic cipher, Vigenère. Students learn how it addresses some of the problems in monoalphabetic ciphers, but also how it is still vulnerable to key length and frequency analysis attacks. To tackle these issues, the module finishes with Enigma, the polyalphabetic cipher used in World War II by Germany. While Enigma addresses the weaknesses of the prior ciphers, students then learn about the problems Enigma suffered from including known-plaintext attacks and key distribution. By subsequently showing students the movie "The Imitation Game", they see how the exploitation of such weaknesses was critical in breaking the Enigma and influencing the course of World War II.

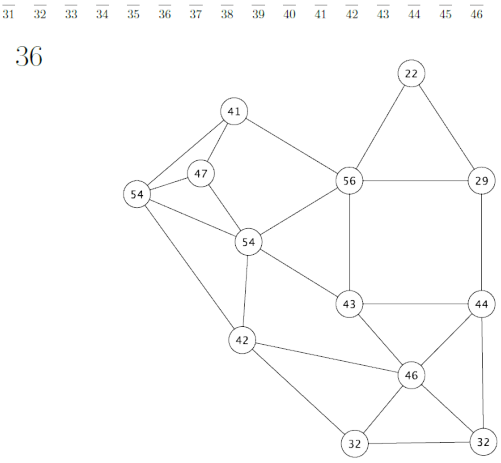


Figure 1: Example collaborative puzzle

In order to solve the key distribution problem, the fourth module then introduces public-key cryptography and the notion of basing encryption on computationally difficult problems. To make the topic accessible, the dominating set problem as described in CS Unplugged is used [5]. By instantiating public-private key pairs via specially-constructed graphs and by using them to encode numbers, students gain an appreciation for how modern ciphers based on public-key cryptography work.

The final module brings students up to current events. While public-key cryptography partially solves the key distribution problem, it is still subject to man-in-the-middle (MITM) attacks in which the retrieval of the public-key is compromised. Using a physical role-playing simulation of a protocol exchange, students learn to appreciate the power that someone inside the network can have in subverting secure communication. By linking this with the GCHQ's FLYING PIG program or the NSA's QUANTUM program for running MITM attacks against a variety of popular sites [6], students are able to link what they've learned with events happening in the present day.

To teach the essential material, the modules employ an alternating lecture and group exercise format where, as soon as a topic is covered, students are tasked with applying what they've learned to solve an in-class, hands-on puzzle. By including collaborative team puzzles, where the overall solution is a combination of each team member's sub-puzzle, we enable horizontal learning in which students who have mastered the topic are able to teach those who have not. For example, Figure 1 shows an ex-

ercise based on the CS Unplugged module on public-key cryptography using the dominating-set problem. In the exercise, each student in the room is given a single character to decrypt in a quote. The quote has 52 characters and each one has been encrypted using the public-key of the recipient. The figure shows character number 36 of the quote. To solve it, each student must brute-force the private-key (the minimum dominating set for the graph), add up the numbers associated with each node in the set, and then lookup the ASCII character associated with the sum. After a student solves their character, they can then help other students as well as work with them to piece together the other 50 characters of the quote.

2.2 CTF challenges

To help students meet the learning objectives of the curriculum, we employ a series of challenges that are organized in a Jeopardy-style CTF. The challenges are scaffolded in a way to allow students to rapidly develop competence and confidence in applying the encoding schemes and cryptographic algorithms covered. Our challenges are inspired by CTFs such as Microcorruption [7] and Natas [8] which allow beginners who may have minimal experience to become proficient in reverse-engineering and web site penetration, respectively. Similar to these CTFs, we employ 24 scaffolded challenges with each one sharing a simple, common gameplay mechanism. By keeping the play uniform, participants can focus on the technical content of each level rather than the game mechanics. To provide continual challenges to students, the difficulty of the technical content incrementally increases to match student proficiency. Levels begin with the simplist of encoding schemes and build their way towards modern cryptography based on mathematically difficult problems.

2.3 Divergent storyline

To better engage students in the learning activity, we design our CTF levels so that they are woven into the plot of a well-known series of books that many young adults of this generation are familiar with: "Divergent" [1]. As levels are solved, students progressively open up parts of the story. This mechanism attempts to provide students with an additional level of engagement and motivation for solving the challenges. We choose "Divergent" as it is extremely popular with the current generation of students. The books have sold more than 30 million copies and the first two books ("Divergent" and "Insurgent") have been adapted into movies which have each grossed close to 300 million dollars. The novels are set in a dystopian future brought about by misuse of technology and raises issues that are highly relevant in an age of immense technological upheaval. It also has a female heroine (Tris) and themes that are compatible with what our discipline needs: highly-trained people with a diverse set of skills and a wide area of expertise (i.e. people who are Divergent).

What makes the "Divergent" series particularly amenable to adaptation to our CTF challenges is that it has a plot in which computer security plays a signifcant role. This theme is explicitly brought out in the short story "The Traitor" which is included in the collection "Four" that was released after the initial trilogy was published. While the initial novel "Divergent" details the events through the eyes of Tris, "The Traitor" details the events through the eyes of the other main character of the book, Tobias (a.k.a. Four). This is of particular interest to us because Four is a bit of a hacker and works to break into computer systems in order to glean information. In the short story, Four employs shoulder surfing, backdoors, trojans, and rootkits to compromise the computer system of his faction's leader (Max) in order to uncover a plot that Max has with a rival faction (Erudite) to annihilate a third faction. The short story details the security training Four receives in getting his position in the computer control room and the techniques he uses to bypass its systems as well as those of the faction's leader.

Using the story as a basis, the CTF challenges are integrated into it using a plot device centered around an electronic diary. In the CTF adaptation, Tris contacts the student participants at the beginning of the week with an urgent message, communicating with them as if they are fellow faction members. She tells them that Four has mysteriously disappeared and that the only clues she has been able to find in rifling through his apartment are a USB key which seems to contain individual entries to a password-protected electronic diary and some undecipherable printouts that might be the passwords to unlock the diary entries. Given the cut-throat nature of the faction, it's likely that Four was paranoid about someone discovering what he was up to and used the cryptography he learned from his computer control room training to protect his diary. Since she knows that the students are going through similar training (i.e. the camp), Tris asks for their help in deciphering the diary and figuring out what Four might have been working on. She also tells them that although Four loved Jumble puzzles, the passwords seem to go beyond those simple letter scrambles.

With this as the setup, each of the printouts contain our scaffolded CTF challenges in which students must apply their knowledge of data encoding and cryptography to solve. The printouts contain encoded clues and information that eventually lead to the passwords that can decrypt individual diary entries. The difficulty of each

(a) Chained cryptographic puzzle

(b) Wooden caesar wheel given to students

Figure 2: Example puzzle

password steadily increases as students progress. Initial entries employ the simplest mechanisms such as ASCII encoding or a barcode while later entries use more difficult ones such as public-key encryption as well as those that chain multiple mechanisms together such as the one shown in Figure 2.

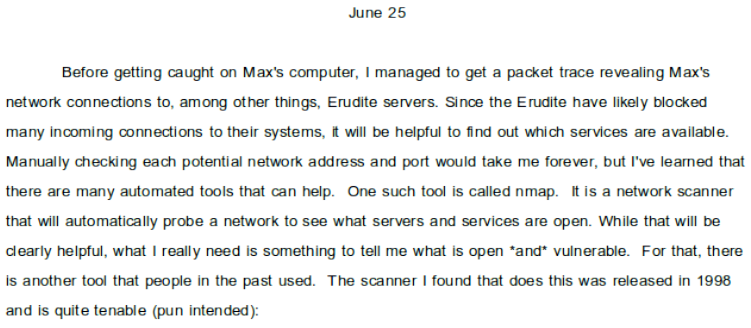
2.4 Security Jeopardy!

While the password-protected diary provides a straightforward structure to each of the CTF levels, it also allows us to address our other main learning objective: teaching students about computer security mechanisms in a context they might remember. From the unlocking of the first diary entry, students find out that there is a security puzzle for them to solve within. Each diary entry is a first-person account from Four of the steps he has taken to compromise the computing systems of the faction's leader and eventually the computer systems of the Erudite. The diary entries initially follow the plot as described in "The Traitor" closely and are effectively like reading a diary of a penetration tester. By describing the issues and technologies involved in computer security using a character and a story that many students connect with, the aim is to invoke an appreciation and a curiosity in security.

The 24 entries of the diary are set within the preceding month and each entry describes a tool or technique that Four has used (or one that has been used against him) in his cat-and-mouse game of trying to infiltrate the computing systems of his adversaries. Ever the paranoid one, rather than spell out exactly the technology and technique he has used, he roughly describes it (as in the Jeopardy! game show) and forces students to figure out what it actually was. Only by solving this puzzle within a puzzle and giving its answer back to Tris, can students complete the entire level. Thus, in addition to giving students hands-on practice with the curriculum topics, each level also requires students to perform research into different aspects of computer security.

In its initial instantiation, the diary covers Four as he:

- Uses a surveillance camera to obtain Max's password and gain initial access to his computer
- Installs a backdoor to maintain access on Max's computer
- Discovers the intrusion detection system protecting Max's system
- Exfiltrates data from Max's system covertly
- Covers his activity in the control room
- Decrypts files from Max's computer that use weak modes of encryption and weak passwords
- Deploys network monitoring to capture information in real-time
- Is caught via use of a fake program
- Attempts to compromise Max's hardened replacement computer
- Initiates a social engineering attack on Max that fails as a result of a password manager
- Launches a session-hijacking attack that fails due to the use of HTTPS and script blocking


June 25

Before getting caught on Max's computer, I managed to get a packet trace revealing Max's network connections to, among other things, Erudite servers. Since the Erudite have likely blocked many incoming connections to their systems, it will be helpful to find out which services are available. Manually checking each potential network address and port would take me forever, but I've learned that there are many automated tools that can help. One such tool is called nmap. It is a network scanner that will automatically probe a network to see what servers and services are open. While that will be clearly helpful, what I really need is something to tell me what is open *and* vulnerable. For that, there is another tool that people in the past used. The scanner I found that does this was released in 1998 and is quite tenable (pun intended):

Figure 3: Example diary entry with Security Jeopardy! challenge

- Performs anonymous reconaissance on Erudite networks to identify weaknesses
- Exploits a range of vulnerabilities to gain access to protected Erudite systems
- Discovers that there is a single air-gapped system that contains everything he is looking for and what is required to compromise it.

With each entry, students uncover and research state-of-the-art security techniques and tools that Four uses to infiltrate systems. Figure 3 shows an example diary entry in which Four employs an industry-standard vulnerability scanning tool to try and find weaknesses in the Erudite network. As the figure shows, after reading the entry, campers can then research and find the actual tool that he has employed.

2.5 Urban race

After developing the storyline at the beginning of the week, the CTF challenges eventually lead the students to a climactic event that is occurring in the present and forces the students into the plot itself. Specifically, in the the last entry, Four reveals that the main Erudite computer system is air-gapped and that he has gone to infiltrate Erudite headquarters in order to break into it. It is at this point that we drop students directly into the plot as full participants. In a timeslot initially labeled as a lecture, we instead pivot to an urban race activity by springing upon students an urgent message from Tris. Tris has gotten a message from Four who has managed to make it just outside of the Erudite control room containing the air-gapped computer. The Erudite have rigged access to the control room in order to ensure that only true Erudite are allowed in. To do so, advanced cryptographic puzzles that only Erudite can solve are given to those that seek access. Moreover, as an added authenticator, solving these puzzles requires specific knowledge of the Erudite campus. Four has just been issued a set of challenges that must be correctly solved within a short period of time, otherwise an alarm will be raised and he will be caught. He needs help solving the challenges quickly in order to to thwart the attack plans.

Tris delivers the students the set of cryptographic clues Four has just been given. They must race to decode the puzzles which then leads them around the Erudite (Portland State) campus searching for answers in order to help the main character of the story in real-time. Modeled after urban races such as CitySolve and Challenge Nation, the clues send teams throughout campus and require team members to communicate virtually with the main character as he submits challenge solutions to authenticate himself. In the current instantiation, communication with Four is done via Twitter direct messages. A "virtual" Four, written as a Twitter bot, provides the illusion of interacting with the actual character. As teams interact with Four, they are given feedback as to which of their answers worked and whether or not the story ends positively. For the race, each team has a separate instance of the story being simulated and the bot keeps track of the progress of individual teams. Successful teams are able to get Four access into the Erudite control room, while unsuccessful teams force Four to take matters into his own hands [1]. While the goal of the race focused mostly on getting all teams to complete it, we also wanted to recognize and reward the team that was able to work together to complete the race first. Towards this end, the virtual "Four" sent the first team to complete the race a final puzzle that led them to a prize, which in this case, was an invitation to dinner that night with a local Turing award winner (Ivan Sutherland). Figure 4 shows an example interaction between the winning team of the Urban Race and the character Four, as embodied by our Twitter bot.

[1] There were no unsuccessful teams in our initial instantiation

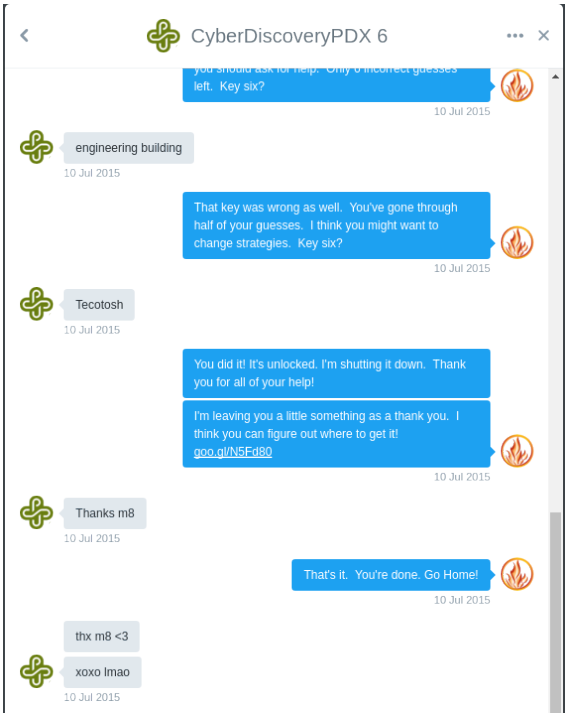


Figure 4: Twitter bot interaction for urban race

Figure 5: Group collaboration during Urban Race

Venue	Date
CyberDiscovery Portland State	7/2015
Portland State New Beginnings	9/2015
Lewis and Clark College	1/2016
Lincoln High School	Spring 2016
Sunset High School	Spring 2016
CyberPDX Portland State	7/2016

Table 1: Camps and classes employing curriculum

2.6 Reflecting on "Flow"

One of the most powerful psychological states humans can achieve is what Mihaly Csikszentmihalyi termed "Flow". The state is characterized as one of single-minded focus on a task and aligns a person's emotions and motivation with the objective at hand. Anecdotally, activities such as CTF competitions engage participants at a level that allows such states to be experienced and are a part of what makes them so popular. Towards this end, the design of our curriculum is explicit in its intent to deliver this experience to students and has design elements that are considered 'flow' triggers [9] for keeping students engaged in the learning activity.

Specifically, with the CTF challenges, there are clear goals to each level, there is immediate feedback as to whether or not a level has been solved, and there is a balance of challenge and skill that is intended to reduce boredom and frustration with challenges whose difficulty steadily increases. Both the CTF and subsequent urban race also bind the solutions of individual challenges to the progression of a story, providing a richer environment for the learning activity.

With the urban race, 'flow' triggers include increasing the risk involved (i.e. the fate of the virtual characters that students have a connection with), providing each group with a shared common goal, and requiring constant communication amongst group members to successfully complete. Moreover, as a team activity, the urban race forces each member of the team to apply the skills they have been taught and encourages members to teach these skills to each other. Figure 5 shows one of the student teams as they work together to complete the race.

3 Results

Our initial instantiation of the curriculum was offered as the cryptography thread of Portland State University's CyberDiscovery camp for rising 10th grade high-school students held in July 2015. In the camp, 9 teams of high-school sophomores (51 students total) completed all of the challenges in both the CTF and the Urban Race. The success of this offering and the interest that it generated led to offerings in several surrounding high-schools and colleges. Table 3 lists the instantiations that have been run. Note that because the Urban Race activity is directly tied to locations on the Portland State campus, only camps and classes located at Portland State employed it.

To evaluate the saliency of the curriculum and CTF, Table 3 lists the results of a survey given to students who completed the curriculum at Lewis and Clark College.

Question	Average rating (1=Not helpful 5=Helpful)
Lecture material for helping learn cryptography	3.82
CTF format for helping learn cryptography	4.45
CTF story for helping learn security concepts and tools	4.18
Overall curriculum in increasing interest in computer security	4.18

Table 2: Student feedback from Lewis and Clark offering

The questions sought their opinion on the lecture and CTF formats for helping learn cryptography, on the CTF format for helping learn about security concepts, and on the overall curriculum for increasing their interest in security. As the table shows, students had generally positive experiences, especially with the CTF format.

We continue to seek partners for additional offerings. To help facilitate this, the curriculum and the source code of the CTF and Urban Race are available for all instructors.

4 Acknowledgments

Several individuals were invaluable in helping create and offer this curriculum. In particular, Christian Duncan developed the ice-cream truck activities and the cryptographic treasure hunt activity that preceded the urban race. Christian also was instrumental in running the Urban Race in its first offering. Jean Gourd designed the Caesar cipher wheel given to students while Gerry Recktenwald had them fabricated. In its initial iteration, Tim Sheard and Lois Delcambre stepped in to teach the entire curriculum. Subsequent offerings were facilitated by Jens Mache (Lewis and Clark College), Richard Tinling and Amelia Kawasaki (Lincoln High School), Jason Galbraith (Sunset High School), and Bryant York (Portland State University, New Beginnings Program).

Finally, this material is supported by the National Science Foundation under Grant No. 1623400. The 2015 CyberDiscovery camp that initially employed this curriculum was supported by the Cyber Innovation Center through the Cyber Discovery Program at the National Integrated Cyber Education and Research Center. The 2016 CyberPDX camp that hosted the current version of the curriculum was supported by the GenCyber Program jointly supported by the National Science Foundation and the National Security Agency. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or any supporting agency.

References

[1] V. Roth, *Divergent Series (Divergent, Insurgent, Allegiant, Four)*, HarperCollins Publishers, 2014.

[2] M. Gladwell, *Outliers: The Story of Success*, Little, Brown, and Company, 2008.

[3] Code.org, "The Hour of Code," `https://hourofcode.com`.

[4] W. Feng, "CyberDiscovery Cryptography Thread," 2015, `https://cyberd.oregonctf.org/`.

[5] CS Education Research Group, "Dominating Sets," `http://csunplugged.org/dominating-sets/`.

[6] "Flying Pig document leak," 2013, `https://www.scribd.com/doc/166819124/mitm-Google`.

[7] "Embedded Security CTF," `http://microcorruption.com/`.

[8] "OverTheWire Wargames," `http://overthewire.org/`.

[9] S. Kotler, *The Rise of Superman: Decoding the Science of Ultimate Human Performance*, New Harvest, 2014.