National Cyber League challenges as a way to teach Cybersecurity

André Agosto Quiñones
Master in Computer Science
Advisor: Dr. Jeffrey Duffany
Electrical & Computer Engineering and Computer Science Department
Polytechnic University of Puerto Rico

Abstract *— Cybersecurity education is a relative new topic in most educational organizations today. This presents a problem to most educational institutions like universities who confront the challenge of teaching cybersecurity topics to students that doesn't have the necessary skills and knowledge to understand these topics. Across United States and other countries different initiatives had emerged to deal with this problem [1] [2]. They range from creating all the needed resources in the university or direct the student to resources outside the university. This project is intended to attack the problem by creating labs that teach the students skills and topics in cybersecurity like the ones presented in the National Cyber League (NCL) competition and later use the NCL platform as a testing ground. The main benefit of our approach is being able to fill those skills and knowledge gaps a student can have by tailoring each lab using the NCL competition as a reference [3].*

Key Terms *— Capture the Flag, Cybersecurity, Education, Educational Material.*

INTRODUCTION

The purpose of this paper is to summarize the process in which I incurred in order to develop the labs for a cybersecurity introductory course and the end product achieved using as guideline material the NCL competition. When the idea of this project was presented to me by Dr. Jeffrey Duffany, there were many questions that came to my mind. But three fundamental questions strike me more than anything else and they become the main starting point to stablish the requirements of the project. These questions were: What am I trying to teach? Whom am I trying to teach? What is the end purpose of this teaching activity? Below I will describe the reason why of these questions and how they were answered. Even when I was not going to be the person that was going to teach the course I was going to create the educational material that was going to be used to teach other students. With no prior experience educating at university level I use as guide my experience in other courses as a student and begin to write down what I remember about each course specially those in Computer Science and Digital Forensics that I took with my mentor. Prior to this project I benefit from a cyber security course in Penetration Testing offered by two MS student peers (Yoshuam Alicea, Steven Bennet) who uses CTF challenges to teach about cyber security and later evaluate the students by solving CTF challenges. My mentor Dr. Jeffrey Duffany also provide me with the idea of creating tutorials that students could use to learn about cybersecurity following the kind of exercise or challenges used in NCL. This way of teaching cybersecurity increases student engagement and leads to more well-developed skills [4] [5]. The combination of all this information and experience had driven me in the process of developing the labs.

What am I Trying to Teach?

As specified by my mentor the main goal of these labs is to provide the students with the knowledge necessary to solve cybersecurity challenges at the time they gain more skills. This must be done in a step by step basis they can follow. This demands the analysis of every section of the NCL competition in a way I could extract the necessary information to instruct the students about how the problem could be solved and what tools they could use. The course it is intended to be introductory which implies that each section will be explained at a superficial level without going too in depth. The labs will cover the following areas which correspond to the NCL challenges section:

1. Cryptography
2. Enumeration and Exploitation
3. Log Analysis
4. Network Traffic Analysis
5. Open Source Intelligence
6. Scanning and Reconnaissance
7. Password Cracking
8. Wireless Access Exploitation

Given the volume of sections to work, each section will count with a brief discussion based on definitions, terms, introduction and purpose, and procedure detailing different examples, and a question section for evaluation of the student. This will be enough to introduce the student to that area and he, being able to solve the challenges included in the labs.

Who am I Trying to Teach?

The labs are aim to students with few or none knowledge in cyber security but with computer science or computer engineering background.

Most of the students who will take the course are supposed to be students with programming experience and already had taken their network course. Below is a list of courses they should take first in order to be successful or have a better comprehension of the material in the labs.

Courses:

1. Programming: Computer Programming 1 and 2
2. Network
3. Unix Introductory course: highly recommended

What is the End Purpose of this Teaching Activity?

The purpose of these labs will be to provide the students with tools and knowledge that serves as a base for continuing learning in cyber security and they can improve along time by participating in the NCL competition and eventually other CTFs competitions.

From the prior analysis I construct a basic format to build each one of the labs.

1. Terms and Definitions
2. Intro and Purpose
3. Materials & Tools
4. Procedures
5. Question section and fill out tables

To provide the student with an environment they all can use; an image of a virtual machine will be provided. That way I can assure that all students will have the same tools at his disposition with no need to lose time in installation or configuration. In the next sections I will discuss the labs content.

Labs Content: Cryptography Lab

The cryptography lab will contain the definitions of the following words in (Terms and Definitions):

1. Plaintext
2. Ciphertext
3. Key
4. Cryptography
5. Encryption
6. Decryption
7. Encryption Algorithm
8. Decryption Algorithm
9. Symmetric Encryption
10. Asymmetric Encryption
11. Transposition
12. Substitution

These definitions are provided because we want to show traditional ciphers as well as modern cipher like the public key encryption. NCL challenges include both traditional and modern ciphers.

One example of a traditional substitution cipher that is discussed is the Caesar shift cipher. This one is explained in words and using a table that shows how the alphabet is shifted. To show how it works I chose it to have a shift of 1. In table 1 you can see how a name (ANDRE) will be encrypted using the Caesar cipher with a shift value of one.

Table 1
Name Shifted by 1 using the Caesar Cipher

A	**N**	**D**	**R**	**E**
B	O	E	S	F

Table 2 shows what happens to the first five letters of the alphabet when is shifted by one. This way we show the student a visual representation on how the cipher works.

Table 2
Beginning of an Alphabet Shifted by 1

A	B	C	D	E
B	C	D	E	F

The lab shows how a transposition cipher works by explaining the route cipher. It uses a textual description and a diagram to show his functioning, just like the Caesar cipher. In figure 1 we can see the representation of the route cipher.

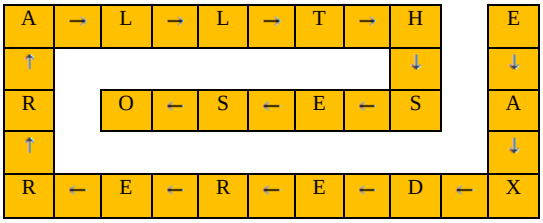


Figure 1
Route Cipher Route

Then we discuss public cryptography by explaining how the RSA works. In this case everything is explained in text and by math formulas.

Then we pass to exercise in which I ask the student to solve similar problems, by their own using the procedures taught in the paper or by utilizing the online tools to apply brute force.

Other exercise includes steganography in which a hidden message or file in hide within another file. Steganography is the technique of hide in plain sight. In this case we give the student a message with a hidden message and in other example an image with another image hidden inside of it.

Enumeration and Exploitation

The enumeration and exploitation lab is one of the most important of all the labs. This lab focus on teaching the student the importance of the analysis by breaking problems or things in all his parts and make sense out of them. In this lab the student analyzes a java script code used for authentication and a paragraph containing a hidden message until he finds information of interest that allows him to manipulate it and achieve the objective.

The definitions provided in the lab are (Terms and Definitions):

1. Enumeration
2. Exploitation
3. Vulnerability
4. Threat
5. Resources
6. Element

The procedure begins by reading the code shown in figure 2. After reading the code the student begins to identify parts or elements that could be of interest like variables, functions, math equations, loops, and structure decisions that can speak to him about the process used to validate the credentials. Every element found is written down in a tabulated format to be analyzed by the student as shown in table 3.

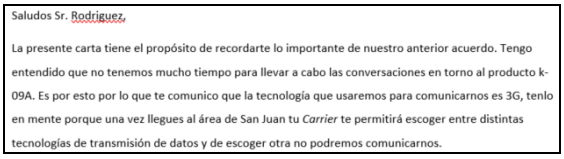
```
<script>
//avaScript get from the webpage to be hacked
// Look's like weak JavaScript auth script :)222222222
    $(".c_submit").click(function(event) {
      event.preventDefault();
      var k = new Array(176,214,205,246,264,255,227,237,242,244,265,270,283);
      var u = $("#cuser").val();
      var p = $("#cpass").val();
      var t = true;

      if(u == "administrator") {
        for(i = 0; i < u.length; i++) {
          if((u.charCodeAt(i) + p.charCodeAt(i) + i * 10) != k[i]) {
            $("#cresponse").html("<div class='alert alert-danger'>Wrong password sorry.</div>");
            t = false;
            break;
          }
        }
      } else {
        $("#cresponse").html("<div class='alert alert-danger'>Wrong password sorry.</div>");
        t = false;
      }
      if(t) {
        if(document.location.href.indexOf("?p=") == -1) {
          document.location = document.location.href + "?p=" + p;
                }
      }
    });
</script>
```

Figure 2
Javascript Authentication Code

Table 3
Variables from the Java-Script Authentication Code

Input	Output	Hardcoded Values
u = holds the user name enter by the user	n/a	k = (set or array of hard coded values) { 176,214,205,246,264,255,227,237,242,244,265,270,283 } //Quantity of hardcoded values:13
p = holds the password enter by the user	n/a	
t = Boolean value		

In figure 3 the student can see the case of a paragraph with a hidden message, the student is confronted with a natural language message which contains tips about the hidden message he wants to share. Then, the student will read the message and write down the part of the message he thinks are important as shown in table 4. Using this information, he must begin to obtain information hidden in the paragraph.


Saludos Sr. Rodriguez,

La presente carta tiene el propósito de recordarte lo importante de nuestro anterior acuerdo. Tengo entendido que no tenemos mucho tiempo para llevar a cabo las conversaciones en torno al producto k-09A. Es por esto por lo que te comunico que la tecnología que usaremos para comunicarnos es 3G, tenlo en mente porque una vez llegues al área de San Juan tu *Carrier* te permitirá escoger entre distintas tecnologías de transmisión de datos y de escoger otra no podremos comunicarnos.

Figure 3
Paragraph with Hidden Message

Table 4
Possible Clues within the Paragraph

Paragraph Number:	**Keyword or Key phrases:**
#1	La presente carta tiene el propósito de **nuestro anterior** acuerdo**.**
	Es por esto por lo que te comunico que la tecnología que usaremos para comunicarnos es **3G**.

Log Analysis

The log analysis lab is intended to teach the students about how to find data in log files that serves to identify an event or a series of events of interest. Due that most log files are ascii files I can use text processing tools to find this data. I can do this by filtering using terms or regular expressions, counting how many times an expression is found, sorting by a term, printing a file and passing it like a stream, and comparing files. The tools are command line oriented and are part of the unix text processing tools.

The definitions provided in the lab are (Terms and Definitions):

1. Log File
2. Logging
3. Analysis
4. Unix Pipes

The Log analysis lab will explain each tool separately, that way the student will understand the individual effect a tool can cause on certain input.

In the figure 4 we shown an example of using the cat command along with the grep command. This way the student can see the output obtained after the execution of the command.

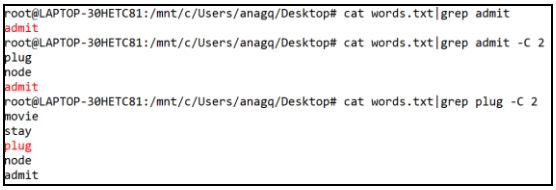


Figure 4
Cat and Grep Command with Context Specified

Every tool used in the lab is presented this way. Then a real log example is used in order to show how all these tools can be used for different reasons on this target.

Network Traffic Analysis

The network traffic analysis is intended to teach the students about how to find data in PCAP files (network traffic packet captured) with Wireshark in order to identify useful information as from what page a malicious code was downloaded or any other activity that could be of interest to the student.

The definitions provided in the lab are (Terms and Definitions):

1. Network
2. Ports
3. Machine addresses
4. IP addresses
5. IPv4
6. IPv6
7. Network protocols
8. Packets

The network traffic analysis lab will show the student how to use Wireshark to capture traffic on a network interface and filter the content of the capture to look for a specific type of incoming connection or a file that was downloaded by using distinct filters and built in extraction tool.

In the following figures we see examples on how the filter works on Wireshark and how the output changes depending on the filter I had selected. In figure 5 we can see the capture of packets before a ping scan had been conduct against a remote machine. Then, in figure 6 we see the capture after the ping scan had occurred. In figure 7 we use a display filter to see those packets produced by the ping scan. The filter consists in specifying that we are interested in packets who used the ICMP (Internet Control Message Protocol) protocol.

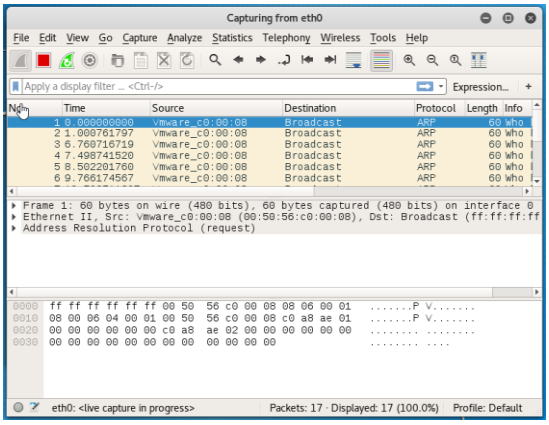


Figure 5

Wireshark before the Guest being Ping w/o Filter

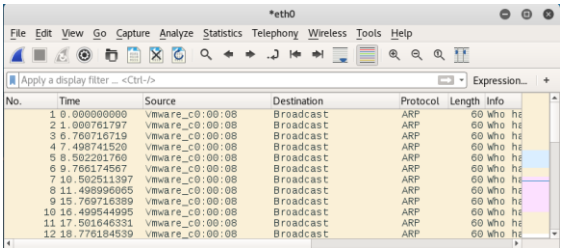


Figure 6

Wireshark after the Guest was Ping w/o Filter

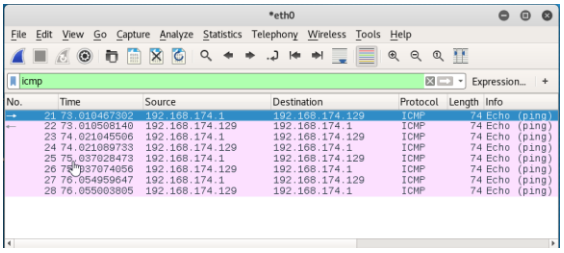


Figure 7

Wireshark after the Guest was Ping with (ICMP) Filter

The idea is that students acquire a basic understanding on how the network communicates with packets and how capturing and filtering these packets by different fields allow us to get precious information about the communication taking place.

Open Source Intelligence

The open source intelligence lab focuses on teaching the students how to look for information using google search engine, to look for computer science and cyber security information which is asked to the student in puzzle's questions. The student will learn to tailor his search by using google's search operators to narrow the results. The lab will concentrate on web search, but it is important to mention that each google search service has a list of allowed search operators.

The definitions provided in the lab are (Terms and Definitions), most of them are intimately related to the search operators:

1. Search field
2. Search operators
3. URL (universal resource locator)
4. Cache
5. File type
6. Site
7. Define
8. Id
9. Info (information)
10. Link
11. Anchor
12. Related
13. Text

The students will look for information on the following topics:

1. Information Security (security concepts, vulnerabilities)
2. Unix & Windows command line utilities
3. Network protocols
4. Programming
5. Publicly available records

Each of these topics will be visited through a single or a set of questions that the student must answer. An example of this is shown below:

What is the common vulnerabilities and exposures code for the Microsoft .NET Framework

Arbitrary Code Execution Vulnerability first published in September 2017?

CVE-2017-8759

If I look only for Microsoft .NET Framework Arbitrary Code Execution Vulnerability I will find various entries with different codes, so I must use the date to narrow my search. I can do this by implementing the following search operators.

intitle:"Microsoft .NET Framework Arbitrary Code Execution Vulnerability" intext:September 2017

This will look for a website that match the "Microsoft .NET Frame…." String in the title and also contains the word and year September 2017 in his content.

The purpose of a question like the one presented before is to make aware the student about the existence of a catalog of known security threats. The catalog is sponsored by the United States Department of Homeland Security (DHS), and threats are divided into two categories: vulnerabilities and exposures.

Scanning and Reconnaissance

The scanning and reconnaissance lab focuses on teaching the student about the process to find out what online services are running on remote machines (computers, servers, etc.) at which he does not have physical access. This lab focus on the "nmap" tool, but for student benefit it explains briefly the purpose of this stage from the penetration tester perspective.

The definitions provided in the lab are (Terms and Definitions):

1. Ports
2. Services
3. Ip address
4. Active Recon (Reconnaissance)
5. Enumeration
6. Scanning

In this lab the student will be introduced to scanning by first being confronted about a set of questions about Nmap functionality and options related to the tool and later he will be asked to scan an online webpage used to check for connectivity and Nmap functionality (http://scanme.nmap.org). Once the scan is complete the student should write down all the services running in the webpage provided with a description about each service running in it.

Password Cracking

The password cracking lab focuses on teaching the student about the proper way in which passwords are stored (hashes), the security this offer and what tools and services can help us when we got the hash of a password and we need to crack it, which means to obtain the plain text password from the hash.

The definitions provided in the lab are (Terms and Definitions):

1. Cryptographic Hash functions
2. Hash
3. Rainbow table
4. Plain text
5. Crack
6. Brute force
7. One-way function
8. Deterministic
9. Message Digest

In this lab the student will be given a set of hashes that will pose as hashes obtained from a credential database or a compromised system and they will have to find out the password or plain text that match with each hash, based on a tip about the plaintext. The student will be introduced to specialized tools like crunch which is useful to generate word lists and "John the ripper". Also, they will learn about the tools included in Kali Linux used to identify a hash and obtained a hash from a string or a file. The hashes will come from different cryptographic hash functions.

Wireless Access Exploitation

The wireless access exploitation lab focuses on teaching the student about the security measures involved in wireless network communication and how these ones can be defeated. In this lab the student will see the process of scanning and capturing packets, but this time with the purpose of

obtaining the credentials to decrypt a secure wireless communication. After presenting general information about this topic the lab introduces the student to the most used aircrack-ng tools and it is ask to find the credential on a wireless packets captured (pcap file).

The definitions provided in the lab are (Terms and Definitions):

1. Packets
2. WEP security
3. WPA V1 and V2 security
4. Brute Force
5. Dictionary Attacks

In order to test the tools the student will use sample capture files included in the aircrack-ng installation package.

FUTURE WORK

As part of the creation of the labs it has been needed to recreate problems which are similar to the ones used in the NCL competition. At first other sources were used because I did not have the exercises that are used in the competence. For that reason, other CTF sites were used to acquire problems and then I modify them. Now I have most of the problems used in the past NCL competition, but they cannot be used as they are because they are copyright.

Therefore, it is necessary to create more problems based on this collection obtained from the NCL competition. This is a job that I already begin but need to continue. Once done they should be added to their respective question sections in the labs.

CONCLUSION

The educational material generated in this project have been used to solve some of the challenges in NCL and other have been generated by solving NCL challenges, this past Spring 2018 competition. There are some topics covered in NCL that could not be covered because of time and resources like the Web Exploitation section, but part of the Web Exploitation section in which we analyze client-side scripts have a lot in common with the enumeration and exploitation section.

The labs are not constructed to be step by step guides to solve CTF challenges instead they are introductory material to gain knowledge and experience in how to solve CTF challenges and by doing that learn about cyber security. Also, the labs provide a way to engage the students in cyber security by exposing them to a more practical experience that will initiate them in a path of continue development by competing with other student peers and cybersecurity professionals. This type of initiative should join another initiatives that will help the educational institution to construct a more dynamic cybersecurity development platform.

REFERENCES

[1] D. L. Burley, “Cybersecurity education, part 1,” ACM Inroads, vol. 5, no. 1, pp. 41–41, Jan. 2014.

[2] D. L. Burley, “Cybersecurity education, part 2,” ACM Inroads, vol. 6, no. 2, pp. 58–58, 2015.

[3] V. Ford, A. Siraj, A. Haynes, and E. Brown, “Capture the Flag Unplugged,” Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education - SIGCSE 17, 2017.

[4] K. Leune and S. J. Petrilli, “Using Capture-the-Flag to Enhance the Effectiveness of Cybersecurity Education,” Proceedings of the 18th Annual Conference on Information Technology Education - SIGITE 17, 2017.

[5] D. H. Tobey, P. Pusey, and D. L. Burley, “Engaging learners in cybersecurity careers,” ACM Inroads, vol. 5, no. 1, pp. 53–56, Jan. 2014.