

Alexander Kerr

# IMPROVING AWARENESS OF ONLINE BROWSING BEHAVIOURS

Learning environment for Phishing attacks

Bachelor's thesis

Bachelor of Engineering

Information Technology

2023



South-Eastern Finland  
University of Applied Sciences

Degree title	Bachelor of Engineering
Author(s)	Alexander Kerr
Thesis title	Improving awareness of online browsing behaviours: Learning environment for Phishing attacks
Commissioned by	-
Year	2023
Pages	37 pages, 15 pages of appendices
Supervisor(s)	Timo Hynninen

## ABSTRACT

It can be challenging to highlight the dangers involved with online browsing from an information security perspective. Following previous research which concluded that providing a more interactive experience aided in raising awareness regarding identifying phishing frauds, this thesis was conducted to research developing a more interactive lab that first-year students were taught during an information security fundamentals course.

The lab was created using virtual machines to provide students with a sandbox area where they can safely learn about phishing techniques and examine the skills involved in some of these attacks.

The results of this research were found by extracting data from student reports. The thematic analysis method was used to evaluate the given responses. The analysis demonstrated that the phishing lab was highly successful in the following areas: Raised awareness of online browsing and potential phishing attempts. Gained knowledge of how these attacks are performed and how to identify them. The ease of these types of attacks, the simplicity involved with creating fake emails.

In addition, an anonymous feedback questionnaire was also used to try and identify the success of the lab. The feedback received here indicated that the students perceived the lab well.

Keywords: cyber security, phishing, security fundamentals, interactivity, raising awareness

**CONTENTS**

- 1 INTRODUCTION.....8
- 2 PHISHING.....6
  - 2.1 Brief history .....6
  - 2.2 Email .....7
  - 2.3 Spear phishing .....9
  - 2.4 Whaling .....10
  - 2.5 Smishing / Vishing.....10
  - 2.6 Angler phishing.....11
  - 2.7 Pharming.....13
- 3 VIRTUALIZATION .....14
  - 3.1 What does the term “VM” mean? .....15
    - 3.1.1 Hypervisor types.....15
  - 3.2 Why use VMs? .....16
    - 3.2.1 XCP-ng.....16
    - 3.2.2 VirtualBox .....17
- 4 IMPLEMENTATION .....18
  - 4.1 Server VM setup.....18
    - 4.1.1 Creating the phishing site .....20
    - 4.1.2 Demonstrating browser warnings .....23
    - 4.1.3 Hosting the web server.....23
  - 4.2 Client VM setup.....24
- 5 ANALYSIS AND FEEDBACK.....27
  - 5.1 Data analysis methods .....28
  - 5.2 Using the thematic analysis method.....28
  - 5.3 Results of questionnaire.....29

6	DISCUSSION .....	31
7	CONCLUSION .....	32
	REFERENCES .....	34

#### APPENDICES

Appendix 1: Lab instructions

Appendix 2: Co-authored article

Appendix 3: Questionnaire used for feedback

## 1 INTRODUCTION

One of the larger issues facing the online community at large is phishing attacks. The practice of phishing has been going on for over 20 years now, the first recorded event of phishing was attributed to a group of users, posing as AOL (America On Line) employees, who were actively spamming other users in an attempt to gain credentials or financial information [1]. I think that we can safely assume that phishing, like the common cold, is here to stay.

To understand modern methods of cyber security awareness training, I searched multiple different sites to see what these types of training normally involved, there was sometimes a breakdown of what these types of courses would attempt to train. The one common theme in these training courses is attempts to raise awareness. The emphasis tends to be on identifying phishing emails and demonstrating what could happen if malicious links are clicked [2-4].

Another challenge is training and campaigns are not effectively engaging with the people involved [5]. These types of campaigns are designed to insinuate fear, it does not aid in raising awareness and oftentimes may prove to be highly ineffective.

Previous research into more interactive methods of cyber training performed by Z. A. Wen et al. has shown that providing a more interactive experience has aided participants' awareness, "What.Hack achieved a 36.7% improvement in players' correctness in identifying incoming phishing emails" [6 p.2].

It is hoped that the use of a more interactive method will let students gain a deeper understanding of how to identify and avoid these types of attacks, and ensure best practices are used to secure data. The skills and practices learned should be transferable to working life.

The challenge at hand is how to engage with students who may have little to no experience in IT (Information Technology) and have them examine their online

browsing behaviours. Everyone should be taught basic awareness and information security fundamentals as soon as possible. It does not matter how much you spend on securing a system from external threats, all it takes to compromise that security is a single employee who does not understand the consequences of their actions [7].

This thesis has the following layout covering several chapters and sub-chapters.

The following is a brief breakdown of each chapter.

- Introduction – Aim of the thesis
- Phishing – Brief history and types of attacks
- Virtualization – Description and uses
- Implementation – The lab creation
- Methods – Methods used to obtain the results
- Results – Where this thesis succeeded and where to improve.

## **2 PHISHING**

The term phishing can encompass several various attacks, all of them however involve a malicious actor who is attempting to fraudulently obtain information/credentials/funds by attempting to impersonate a legitimate point of contact within a firm or organization. This section will discuss phishing and the various forms of attacks that are commonly attempted daily.

### **2.1 Brief history**

There are theories to be found which would point at phishing attacks being the evolved electronic form of the original “Spanish Prisoner” letter, a scam perpetrated in the late nineteenth century [8]. It could be debated if this is true or not, regardless, it is a widespread issue that is constantly evolving with modern-day technology. To highlight the growth of these attacks we can examine the quarterly findings from the Anti-Phishing Working Group (APWG) for the first quarter of 2012 and the third quarter of 2022.

In the summary table for Quarter 1 of 2012, the number of detected attacks is 164,023 [9 p.3]. If we compare that to Quarter 3 of 2022 [10 p.3] which had a

staggering 1,025,968 detected attacks, we can see that phishing attacks are a serious online threat that has been steadily growing over time.

## **2.2 Email**

While most phishing attacks are generally attempted using email when we say the type of phishing attack was email, we are saying that these attacks are arbitrarily sent, the person or people behind the email are, to coin a phrase, “casting out a net” in the hopes of randomly tricking someone. These types of attacks usually claim to be from some genuine institution such as banks or online retailers. They will attempt to persuade you to click on the link provided. Users may also be prompted to download malicious software in attempts to either gain access to the computer, or the data stored on it, the link may re-direct the user to a site which is almost identical to the real service it claims to be. The link may also (Figure 1) attempt to open an email to establish a dialogue with the hopes of convincing the target to reveal credentials.

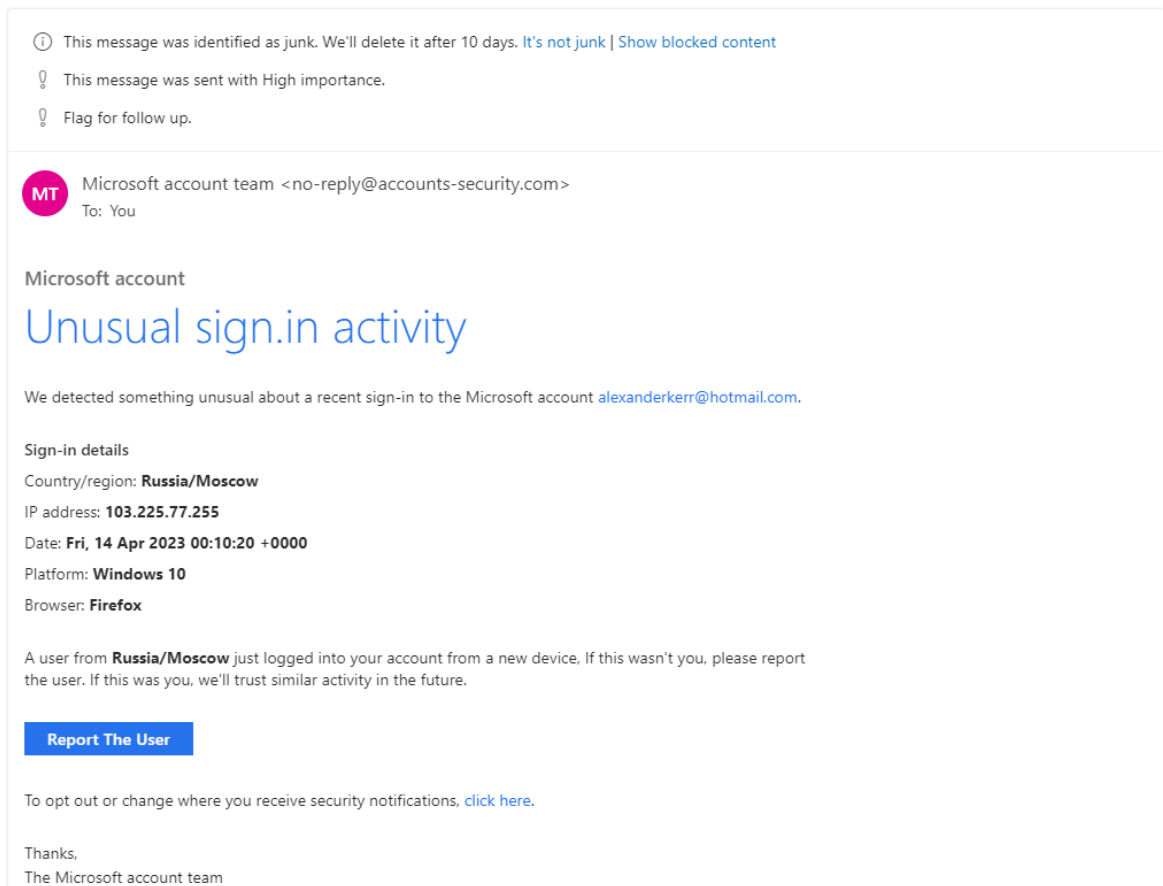


Figure 1: Screen capture of a genuine phishing email from my own personal email

In this illustrated case, if I hover over the suggested “Report The User” button (Figure 2) I can see what action will be performed upon clicking this link.

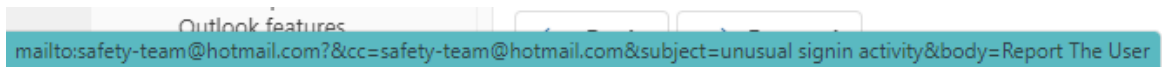


Figure 2: Create an email with auto-filled fields to a user with a Hotmail account.

Clicking on this button would prompt my email client to create a new email to begin a dialogue with someone pretending to work for Microsoft. The email has already been identified as junk mail by Microsoft, but this is one of several hundred phishing emails I will likely receive this year, the majority are normally flagged as junk email, but several attempts will still make it to my inbox.

Typically, this type of phishing is attempting credential harvesting, prompting the user to download malicious software, and advance-fee fraud. The premise of



advance-fee attacks [11] is that a malicious actor portrays themselves as having vast amounts of wealth, which is currently inaccessible. They would claim to require financial assistance from the recipient of the email, with the promise of greater returns later.

The term “419”, [12] the law code used in Nigeria to cover Advance-Fee fraud, is now commonly used in online media to describe email phishing attempts in various online media.

### **2.3 Spear phishing**

This type of phishing attack is a more targeted effort against organizations or individuals. The malicious actor/s involved in this type of campaign have to put in a lot more effort to create these types of attacks since it requires more knowledge of their intended victims. Using social engineering techniques, which will not be discussed here as it falls outside the thesis scope, the attacker will attempt to find out information of whom the company is dealing with, or research individuals through social media posts, they may even attempt to gain access to smaller companies who deal with the target in order to find out names of contacts and check email communications in an effort to create a more convincing campaign.

The aim of this attack according to Williams’ et al. [13 p.1] is “... to persuade employees to click on malicious links, download malicious attachments or transfer organizational funds or other sensitive information”. These types of attacks, if successful, can generate large amounts of revenue for the fraudsters. A classic example of this type of attack was a campaign launched against Google and Facebook [14], the campaign managed to take approximately \$100 million USD from two of the largest companies in the world. This campaign was primarily successful due to setting up fake firms and invoices which were sent to lower-level staff who, perhaps, should have been trained better to check that the invoices were for genuine services provided.

## **2.4 Whaling**

It can be hard to differentiate the difference between spear phishing and whaling, they are both targeted attacks at individuals and organizations. The main difference to be found in whaling attacks is that they are targeting higher level staff such as CEOs and executive level staff.

The methods are similar to spear phishing, the attackers are relying on the fact that higher-level staff may not know all the clients in place and most likely will have no idea about any systems in place for lower-level staff which would prevent a successful spear phishing campaign.

The real threat with whaling attacks is that they are also used for other reasons than just immediate financial gain, higher level staff credentials would make corporate espionage easier, and if an attacker can gain access to the network via executive or CEO credentials, they could cause untold damage both to the network and the business reputation before being discovered.

Successful whaling attacks [15] can be highly damaging to personnel within your company, or financially devastating if not detected in time. It however illustrates that all employees within a company should have mandatory training in these areas to prevent phishing attacks from being successful.

## **2.5 Smishing / Vishing**

Phishing attacks are not limited to email, this section will cover the two types of attacks that are carried out over telecommunication systems.

Smishing, the act of sending an SMS, Short Messaging Service, or as it is commonly known, a text message. Just like email attacks, most of these attacks are randomly sent to mobile phone numbers and usually re-direct the user to an official-looking site. In December of 2021, I almost fell for one of these schemes. I had ordered a gift from the UK and due to my lack of Finnish grammar skills,

received a text claiming that a parcel from the UK required a minor payment to clear customs. Without thinking, I opened the link provided in the text message and entered my information. After submitting my information, I quickly realized that the page looked odd. Checking the website address I had been directed to, I quickly phoned up my bank and informed them of my mistake. Fortunately, the bank was able to cancel my previous card before any funds had been removed, and promptly sent out a new card. This type of attack is currently gaining more traction here in Finland currently [16] [17] [18].

Vishing is the act of phoning an organization to gain credentials or funds, again, this is something I have previous knowledge of. At one point I was receiving at least one telephone call to my house phone, the callers would claim to be from Microsoft or some cyber security firm and they would mention I had a virus on my computer that had to be fixed. They would request that I download some software from a website in order for them to gain remote access to my computer so they could remove the virus. For me it was obvious that this was a scam and I usually ended the call with "I have Linux installed, how can Microsoft help". The problem is that once the fraudsters obtain any response from a called number, they will persistently call in the hopes of finding another household member who will fall for the scam, or they will change their tactics by claiming you had won some prize draw and they just required some personal details from you.

## **2.6 Angler phishing**

This type of attack is carried out over social media platforms such as Facebook and YouTube. On posts that attract a lot of comments, users will create fake accounts which claim to be linked to the original post. The malicious actor will then reply to comments claiming that the original poster has won some sort of prize, or they will target disgruntled complainants pretending to be a customer service representative. The end goal here is to gain credentials.

For clear examples, we need only look to Twitter with services such as Dominos Pizza commonly being targeted by fraudsters [19].

Angler phishing however seems to be on the decline as social media platforms implement better security checks on accounts, this makes it harder for fraudsters to create fake accounts to attempt fraud. The rise in Angler phishing attacks coincided with the Covid-19 pandemic which seen people stuck at home, spending a lot more time on social media platforms. Now that the situation has somewhat normalized, we can compare the APWG report for Q1 of 2021 [20 p.5] (Figure 3) vs the Q3 of 2022 report [10 p.5] (Figure 4). We can see that during the pandemic Angler phishing had 23.6% of the detected attacks.

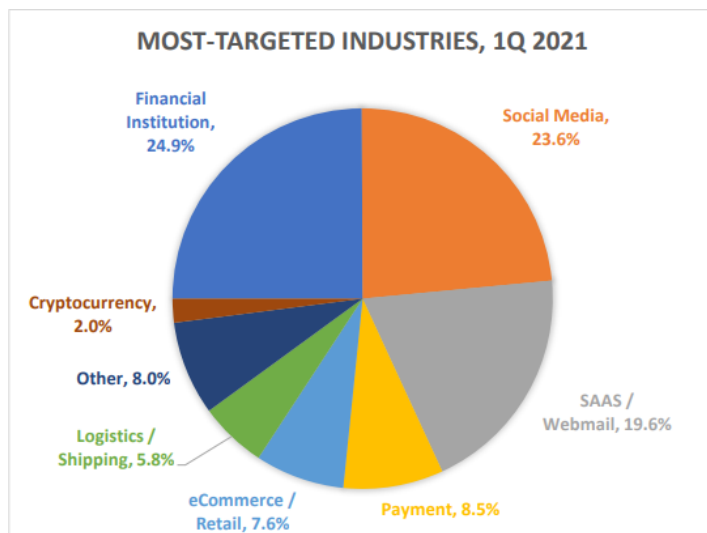


Figure 3: APWG pie chart Q1 2021 [19 p.5]

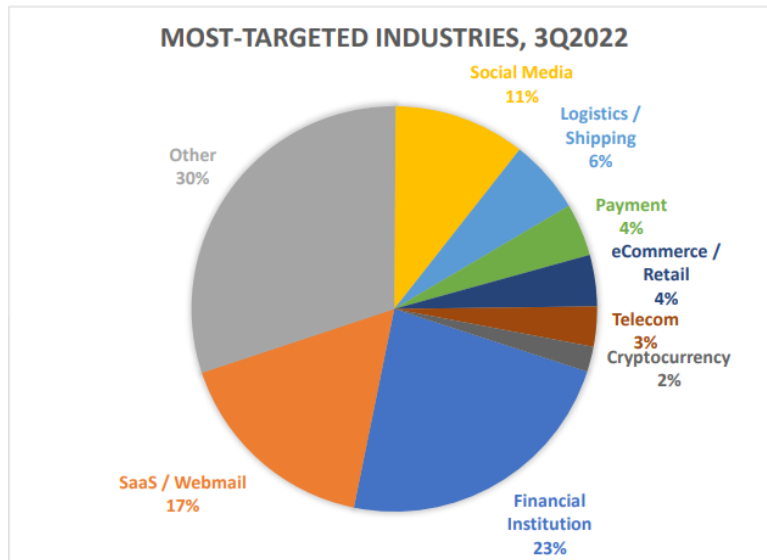


Figure 4: APWG pie chart Q3 2022 [9 p.5]

Whereas late in 2022, Angler phishing accounted for 11% of the detected attacks. The charts clearly demonstrate that fraudsters are constantly evolving and changing their methods in efforts to create new schemes which people are not aware of in order to defraud more victims before their methods are broadcast to the general public.

## 2.7 Pharming

Pharming requires access to either the hosts file of a target machine or to the DNS (Domain Name System) server supplying services to the target network. Through altering the hosts file or performing a DNS poisoning attack, the malicious actor can redirect traffic destined for a genuine site to a fake version that they have control of. The danger posed by this attack is that a user may reveal their credentials to the attackers, who either redirect those credentials to the genuine site which would log the user in, or perhaps just redirect them to an incorrect username/password page. It is not unusual to occasionally type credentials incorrectly so the target may not even realise they have just exposed their credentials.

There are some who say that Pharming is not phishing [21], the argument being that pharming involves gaining access to the local machine or network in order to

alter the DNS settings through either a DNS poison attack or by altering the hosts file of target machines, the hosts file is a record kept by your OS (Operating System) and can be altered to link a specific URL (Unique Resource Locator) to a pre-specified IP (Internet Protocol) address, the hosts file can then perform DNS functions on the local machine rather than querying a remote DNS server.

I would argue however that pharming is still a phishing attack, other forms of phishing might involve gaining access to other networks in order to examine communications between people, once these communications have been analysed, the actor can then attempt communication with their intended target in an effort to convince them that they are a legitimate point of contact. To simplify phishing as nothing more than social engineering coupled with an attack vector is an over simplification. We only need to examine business email compromise (BEC) [22] to see that not all targeted attacks are based purely on social engineering techniques.

The difference between spear phishing/whaling and pharming is that pharming requires access to the target networks DNS service, or to the target machines hosts file. As described here [23] [24] “Pharming is ‘phishing without a lure.’”, which I believe gives credence to pharming being a type of phishing attack. A key pointer that you are being subjected to a pharming attack would be typing the URL of a site that you use regularly and noticing that it has an HTTP (HyperText Transfer Protocol), a protocol which typically transmits data such as usernames and passwords unencrypted, version of a site you would normally assume to be secure with HTTPS (HyperText Transfer Protocol Secure), a protocol that supports encrypted transfer of data such as usernames and passwords.

### **3 VIRTUALIZATION**

This section of the thesis will go into more detail regarding Virtualization and nested virtualization. This section will also discuss the benefits of using virtual environments, and the software used to create the VMs used in the lab.

### **3.1 What does the term “VM” mean?**

When this thesis refers to a VM (Virtual Machine), it refers to software-based emulation of hardware. In other words, using the available resources of the host pc, to emulate other services. While a virtual machine can be created on a host machine, the virtual machine must use fewer resources than the host has available. The host machine needs resources to run and emulate the virtual machine.

VMs can be ran as type 1 or type 2 VMs. I will briefly describe the differences between the two types of virtualization.

#### **3.1.1 Hypervisor types**

Type 1 virtualization refers to VMs that are running directly on the hardware being used. This means that all the resources and hardware are directly available to the VMs implemented. Examples of type 1 hypervisors include Microsoft Hyper-V, XCP-ng, and ESXi [25] [26] [27].

Type 2 virtualization refers to VMs that are being run on a host OS. A typical Windows 11 installation, using minimum recommended requirements, should be run with 2 CPU (Central Processing Unit) cores and 4GB of RAM (Random Access Memory). This means any remaining resources available on the host can be allocated to a single VM or multiple VMs to run on the host machine.

This also means that VMs created as type 2 need to rely on the host OS to communicate with hardware such as memory and the CPU. Type 2 VMs will have a little more latency during use, while noticeable, they will not impact too much on the general running of the VM.

Examples of type 2 hypervisors include VMware player and Oracle VirtualBox [28] [29].

### **3.2 Why use VMs?**

There are several reasons for using VMs [30] [31] [32], however our use of VMs in the classroom is very simple.

Studies in IT often require students to demonstrate practical skills, or to experiment with software which could be used maliciously. Using VMs students can work through various scenarios presented to them safely with no risk to the real hardware and software they are working on.

If students mistakenly or even intentionally manage to damage the OS, they can quickly recreate the VM using the base image that was supplied, within minutes they would have a working VM minus any changes they had made to the previous VM.

If the student makes the same mistake on classroom hardware, the OS will then have to be re-installed along with all the required applications, this will mean classroom hardware that is unavailable for use until it has been restored to its previous working state.

In security fundamentals, students are exposed to software and techniques that would take too much time to configure on each individual computer, then there is the time required to reset the computers configuration and remove any software that students had been working with. Placing all of this in a configurable VM, we can limit the use of the software and techniques being used to the virtual machine. This gives students the benefit of being able to revisit any previous exercises done if they feel that they want to increase their understanding, or if they simply want to refresh their memory of the steps involved.

#### **3.2.1 XCP-ng**

For hosting the web server, we are using XCP-ng to host the VM. As has been described above, VMs hosted on XCP-ng are of type 1, this means the VMs can



directly interact with the hardware and there is no need to rely on an underlying OS to provide intercommunication.

This provides us with the benefit of only using the resources required for the VM. Normally if we just straight installed the OS to the server, all resources are then tied to the installed software. Using XCP-ng means we could create multiple different VMs that could better utilize the resources by sharing them amongst themselves rather than tying all resources to a single OS.

### **3.2.2 VirtualBox**

For students to use VMs we need to consider that they are going to be using classroom hardware or their own laptops with various OS installed, such as Windows 10 or Windows 11, Mac OS, or some Linux variant.

As well as hosting VMs, VirtualBox can be used to create them also, allowing the user to specify the OS, CPU cores, memory, disk space, as well as specifying if the disk space is fixed or dynamic.

A fixed disk space means that if a VM has a total of 100 GB (Gigabytes) assigned to it for storage, then 100 GB of disk space will automatically be reserved from the host machine, making it unavailable to the host OS. If the host machine does not have the required storage free, the VM will be unable to run.

Conversely, if a VM has a dynamic disk space of 100 GB, the host OS will only reserve disk space for what is used, e.g. a typical Linux desktop install with some additional software will use around 10 GB of space. If you only have 40 GB of free disk space left, the VM will run on the host machine, however users would then need to keep checking to ensure the VM disk does not grow beyond the capacity of the available resources.

## 4 IMPLEMENTATION

This section of the thesis will give a description of how the server and client VMs were configured and setup. I have created a simplified topology diagram below (figure 5). The normal flow of events is that when a user types a URL to a web browser, the browser checks if a DNS server is available to translate the URL to an IP address it can visit. This chain of events can be changed though by editing the hosts file of an OS. By default, the hosts file contains no entries, if the hosts file is edited to link a URL to an IP address, the hosts file would be used for the entries listed, before checking DNS services.

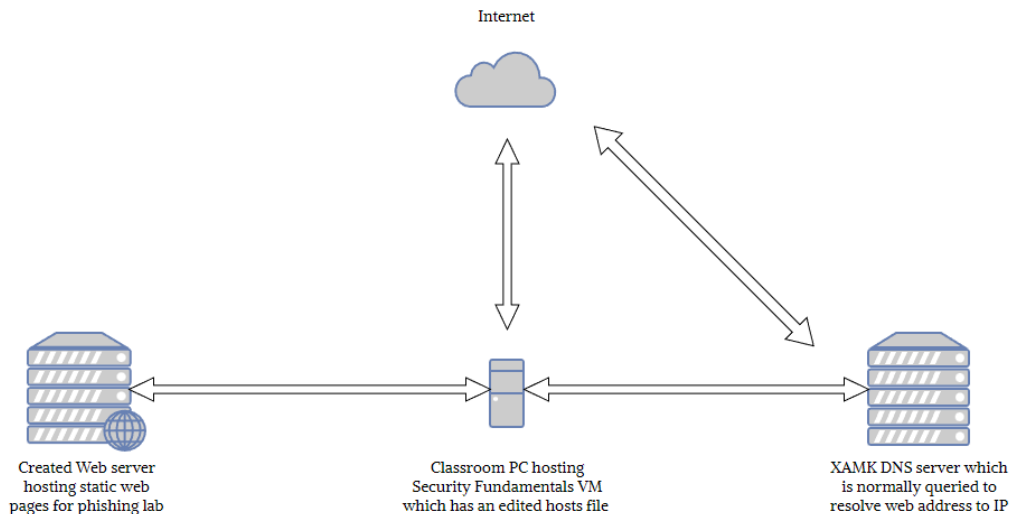


Figure 5: Simplified network diagram for illustrative purposes

### 4.1 Server VM setup

The primary objective for this server was to host a static web page for use in a security fundamentals course exercise regarding phishing.

Since I would be installing web hosting software which would require testing as well as hosting static webpages, I decided to install Linux Ubuntu 22.04 desktop version to the VM. Depending on the version of Linux being installed, the installation process may differ, however there are plenty of guides available with a quick prompt in a search engine on how to install Ubuntu desktop [33]. As an

operating system, Ubuntu is less resource intensive than Windows, it is also open-source and free to use for everyone which removes the requirement of having a product key for the OS. Having created multiple VMs using various Linux or Windows versions I can speak from experience when I say it is a much faster and simpler process to install Linux vs Windows. Using the desktop version I can quickly test the page serving functions within the VM and would not have to create a configuration that would allow the host machine to access the web server to check that the content was being served correctly. This would also make it easier to just focus on the task at hand rather than swap between the VM and host machine just to check that any changes I was implementing were working as intended.

With the OS chosen, I quickly went to work creating the VM in VirtualBox. I chose a single CPU core since the server would be doing nothing more than serving static web pages, in order to install Ubuntu desktop, the VM requires a minimum of 4 GB memory, for disk space, I used 20 GB of dynamic storage, this should be plenty of space to install the required OS, software, and also leave space available for updating the server to keep it secure.

For web serving, I could have used either Apache or Nginx, (pronounced Engine X). I chose Nginx since it is perfect for hosting static pages and is a much lighter software to both install and run compared to Apache [34]. This means we can minimize the resources required to run the server without losing any performance of the services provided.

With Ubuntu desktop now installed to the VM I enabled the inbuilt firewall provided with Linux, ufw or Uncomplicated Firewall to give it its full name, the main account was given an appropriately long password along with the root account.

Once the appropriate security measures have been implemented, it is now time to examine how to install and configure Nginx. I had previously experimented with

implementing some basic web pages in Apache however I had never used Nginx, after searching through some results from my search engine prompt, I quickly found the perfect guide [35] which details the install procedure, allowing the web server to operate through the firewall, and how to setup server blocks. Server blocks allow the web server to host several different web sites if required, each contained within its own server block.

#### 4.1.1 Creating the phishing site

The next step was to implement the web page. All I had to do was pick a site and check if I could see the source code for the page, if I was able to view the source, I could then check if I could see the CSS (Cascading Style Sheets) files which are required for styling how the content appears on the web page, the final check is whether I can download the images also present on the page.

Since we are dealing with XAMK students who regularly use the Learn platform (Figure 6), I decided this might be an appropriate demonstration to use since it should be a familiar page.

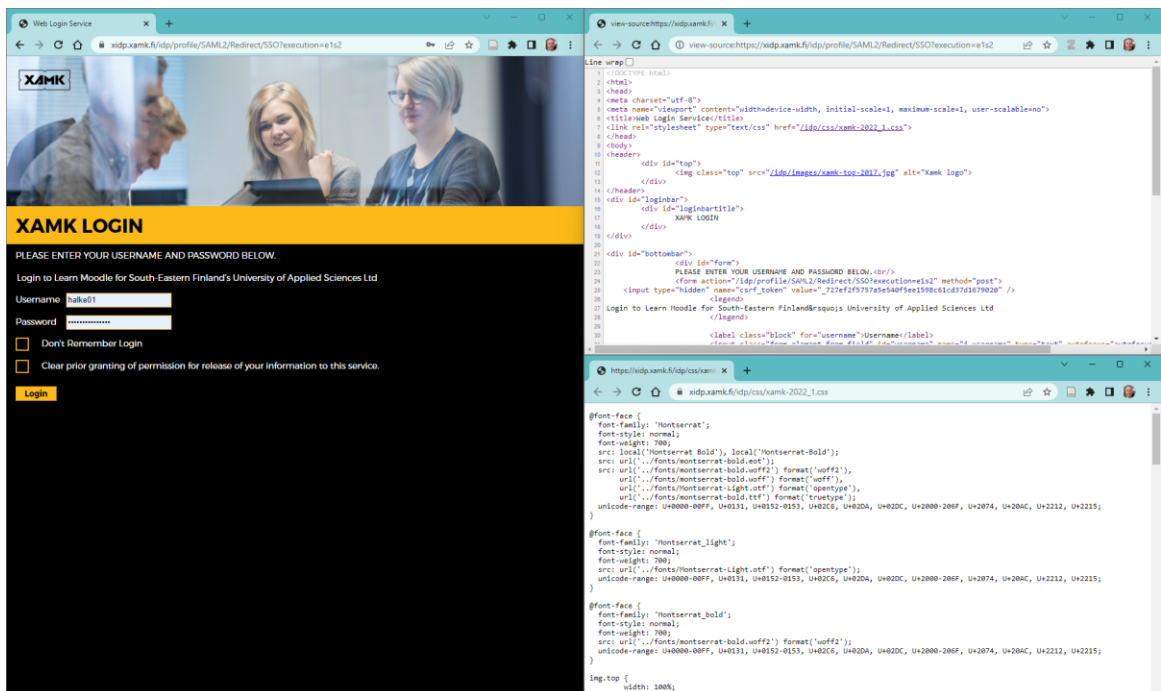


Figure 6: Checking if I can replicate the page with minimum effort

Opening a blank text document in Ubuntu, I copied over the Hypertext Markup Language (HTML) to blank text file, this was placed in the root directory of the server block that was created in the previous section. Naming the file as 'Index.html' makes this file the main file that will load to a web browser when the server block is visited. The CSS (Cascading Style Sheet) was copied over to a second file and the final step was to download any images that were present in the page, these were then placed inside a folder of the root directory and appropriately referenced in the index file.

Using the VM web browser I could check that all elements of the page were present and that there were no large differences to how the genuine learn page looked (figure 7). The only real difference on how the page is displaying is the fonts used, since this is a learning resource, I decided to leave the fonts as they are, there should be some visual clues other than the URL that will make students question the authenticity of the site.

The only question was what to do with the page, as a stand-alone page, you can examine the differences to the genuine page, for example, the different font usage, that it is not the initial page you are presented with if you go to the learn platform, that the protocol being used is HTTP and not HTTPS.

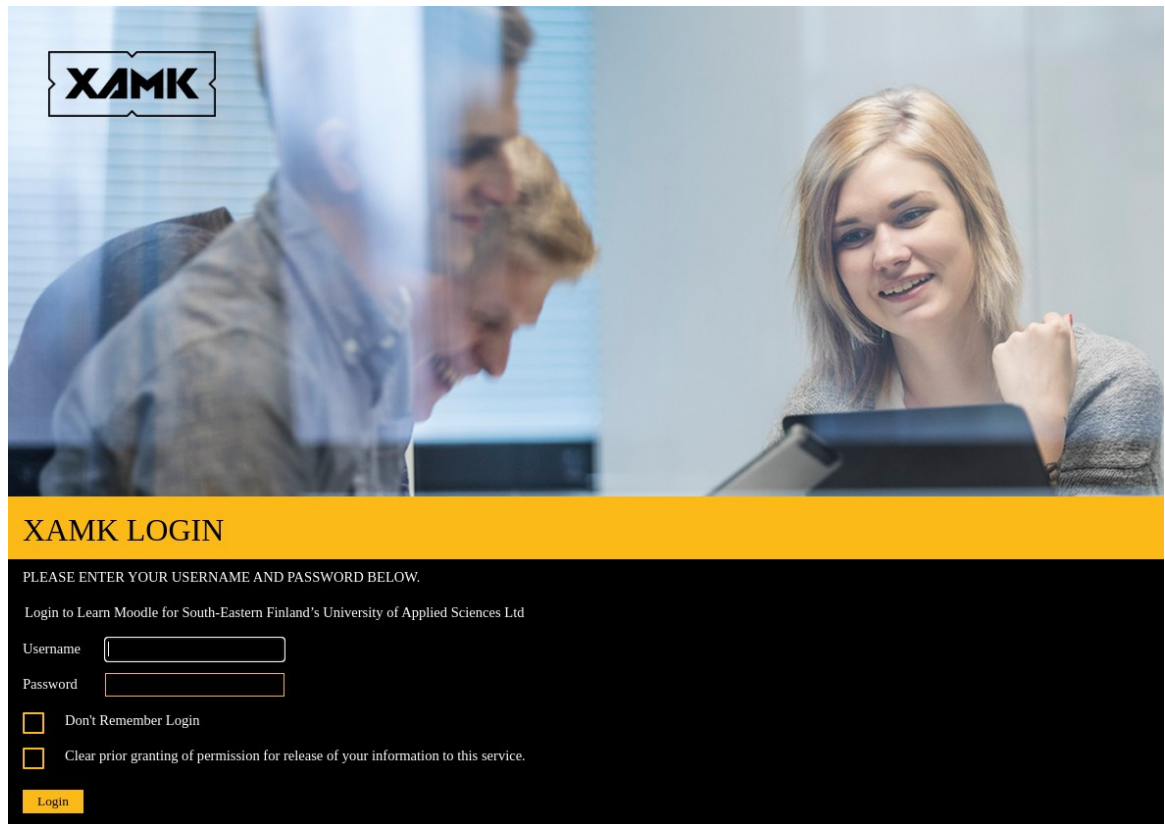


Figure 7: Checking the page displays correctly

Submitting any credentials will just present an error screen which is not very useful. I decided to create a new page to display after the login is clicked. Students should be able to recognise this page is not genuine, however anyone quickly going through the exercise should have a warning that they have fallen for a phishing scam.

The new page I have created (Figure 8) should be sufficient to alert students that they should be paying more attention to detail.



Figure 8: The page displayed after the Login button is clicked

With the new page created, I again had to edit the HTML of the index file. Previously when login was clicked, the page would attempt to post the username and password field values, this action was changed so that instead of attempting to send these values anywhere, they would just be ignored, and the new warning page would be displayed instead. If a student did attempt to login with a genuine username and password combination, those values would not be logged or sent anywhere which would ensure that no one would discover any genuine credentials.

#### **4.1.2 Demonstrating browser warnings**

In order to demonstrate web browser warnings I created and installed a self-signed certificate [36]. Since these certificates should be issued by a trusted authority, creating your own self-signed certificate to authenticate the security of a secure site using HTTPS should prompt web browsers to generate a warning that the certificate in use has not been signed by a trusted authority.

After the web site was tested on the VM, I then checked with the host machine web browsers that the web site was working from an external perspective. Satisfied that my web server was correctly working, my next step was to host it in the local domain.

#### **4.1.3 Hosting the web server**

As was previously mentioned, I want to minimize the impact the VM would have on the available resources we have. Since it is operating with a single CPU and the disk space is dynamic, the only other option I have available is RAM. Modern Ubuntu desktop installations require 4 GB memory in order to install, however, once it is installed, it is possible to run the VM with only 2 GB memory. This will make it run a little less efficiently though since that memory will struggle to provide a smooth graphical user interface (GUI) experience, and deal with any

other background tasks it is supposed to be performing, in my case that would be running a web server.

The simplest method is to disable the desktop and have the OS boot straight to the command line. With no GUI to load and only a web server to run, the Ubuntu install will work great with 2 GB of memory.

Another option is to uninstall and remove all the desktop elements, this will still boot Ubuntu to command line, but will remove the GUI from the system. Not a step that should be taken lightly as it may not be as simple to re-install the desktop. The only reason I had a GUI in the first place was to quickly check my web site was performing as I hoped, without having to rely on other machines. Having now tested my website actually works from the host machine I made the decision to remove the GUI and all connected packages.

The final step is to host the VM on XCP-ng. Using Xen Orchestra (figure 8), which is the web management system for creating or importing VMs, I just need to import the VM, add a network interface, and ask my colleague for a static IP address that the classroom computers can access and assign it to the VM.

The web server is now up and running over the network, I can now visit my replica site by simply typing in the IP address assigned to the VM (figure 9) and it will load the unsecure HTTP version of the site by default.

## **4.2 Client VM setup**

The operating system I chose to use for the client VM was again Linux, as has previously been mentioned, it is quick to install, does not require a license or product key, and is free for anyone to use. Another consideration for using Linux, the majority of previous students have experience using Windows, due to the nature of this lab which will involve altering system files, it is better that we are not teaching students how to do this in Windows, which all the classroom computers have installed.



Since this course will involve freshman students and it is also one of the first courses they will have, it should be expected that the majority of the students will have little or no experience with using Linux or Nginx. It would take up quite a bit of time for students to install and configure Nginx, I asked a colleague to test out my Nginx guide for installing and creating a basic website, even although he is quite proficient in Linux he ran in to one or two issues so the decision was made to have Nginx pre-installed, a server block configured [35], and to add a static web page (Figure 9). Students would then just have to replicate a website within the same server block by editing the existing files and adding the required images and CSS, the instructions for which would be given in the exercise.

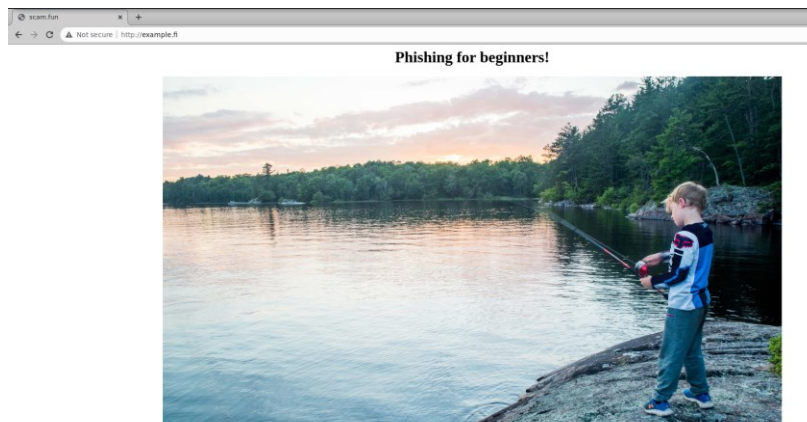
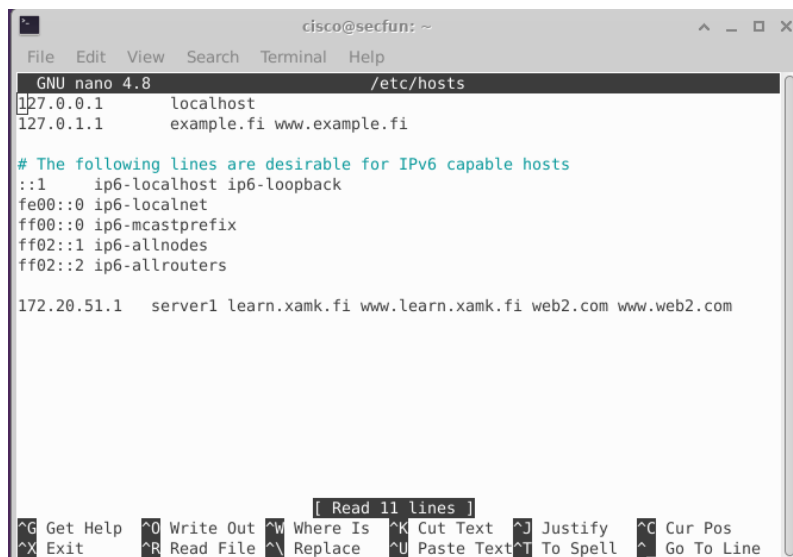


Figure 9: Screenshot of the static page students must edit

The client VM has other requirements such as providing students with the superuser account access, which will allow them to alter the hosts file in Linux, the VM also requires a couple of different browsers to be installed so that students can compare the differences in how the browser will warn the user if the site they are visiting is using HTTP, or if the HTTPS version of the site has a trusted certificate. The VM in question also requires other software installations and user account creation which will be utilised by other exercises in the security fundamentals course, however, these are outside the scope of this thesis and have no relevance to the subject at hand.

The hosts file has to also be edited at this point (Figure 10), we have to point to the locally hosted web site of example.fi on a local host address. The file should also be edited to point at the server VM. The server VM is currently hosted on the IP address 172.20.51.1, any term which has learn.xamk.fi entered in to the address bar of the browser will automatically redirect to our web servers VM. This included web search address, however these would redirect to the HTTPS version of the site which is not the desired outcome when first accessing the page.



```

cisco@secfun: ~
File Edit View Search Terminal Help
GNU nano 4.8 /etc/hosts
127.0.0.1 localhost
127.0.1.1 example.fi www.example.fi

# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters

172.20.51.1 server1 learn.xamk.fi www.learn.xamk.fi web2.com www.web2.com
  
```

Figure 10: Editing the hosts file to redirect web browsers to the phishing page

With the hosts file saved the configuration of the client VM should now be complete. The lab instructions for this exercise can be found later in this document as Appendix 1. Although a brief description of the tasks will follow below.

Students were asked to use the VM web browsers to visit the learn site and attempt to login, students are also asked to examine how 2 different web browsers will attempt to warn the user that the site is not a secure site, the next task is to examine the warning page presented in all browsers when attempting to visit an HTTPS site that is using a self-signed certificate.

Students are then instructed to examine the hosts file so that they can see how the browser is being redirected to a specific IP address if the term learn.xamk.fi is used in the address bar of a web browser. They are then instructed how to remove this redirect and test that the web browsers are now visiting the correct version of the learn site.

Students are then shown how simple it is to replicate a web site that could be used to harvest credentials. Throughout the exercise students are presented with questions that they will have to research online if they do not have personal knowledge of subjects such as HTML or how the hosts file can replace DNS, a service normally supplied externally which will translate a web address to the IP of the web page you are attempting to access.

## **5 ANALYSIS AND FEEDBACK**

To determine how successful this lab was, 2 methods of gathering data were used.

The first set of data gathered came from a reflection report, the task was to choose 3 areas they had learned about and was broken down into a four-part guide as to what the students should be writing about.

- What thing did you learn?
- Why did you choose this thing?
- Why is this thing important?
- How will/can you make use of this knowledge now and in the future?

The second process was an optional feedback questionnaire specifically related to the lab in question. The questions and feedback can be found in Appendix 3 in this document, unfortunately, only twelve out of a potential twenty-six students responded to the questionnaire.

## **5.1 Data analysis methods**

When analysing data, it is important to note that there are several methods available. The first step to choosing a method is to define the data as either quantitative or qualitative. Numerical data being classed as quantitative vs qualitative which is non-numerical in nature.

Once the data has been confirmed as qualitative there are several methods that can be applied to the data. This section of the thesis will examine some of the methods available with brief descriptors.

**Text Analysis:** This method takes large sets of textual data and sort the text into smaller useful units of data. This is normally done by employing text analysis software which has been trained to analyse text and extract the relevant information.

**Content analysis:** This method is used to quantify data, looking for specific words, phrases, concepts, or themes and how many times they are mentioned within the qualitative data.

**Thematic analysis:** This method is used to categorise or codify words and phrases into keywords which can then be assigned a deeper contextual meaning in relation to the subject the responders are writing about.

## **5.2 Using the thematic analysis method**

In order to effectively use the thematic method, the text was examined for any content which would relate to phishing and/or email spoofing, the process of sending fraudulent emails which appear to originate from a different email address. Each responder was placed in an excel sheet, if any relevant text is found in relation to phishing or spoofing, the appropriate text or texts found in the report were then placed in a corresponding cell for further analysis later.

Once the text extracts have been recorded, the next step is to try and codify these statements and group them under keywords. Once keywords have been identified, a clearer description can be ascribed to the keyword and the number of related statements proffered can then be enumerated to find how many instances of these keywords were mentioned (Figure 11).

TABLE II  
AXIAL CODES GENERATED DURING THE THEMATIC ANALYSIS PROCESS

#	Axial code	Number of observations	Detailed description
1	Awareness	16	Descriptions of increased awareness of online security.
2	Critical thinking	2	Descriptions of increased critical thinking about messages on the web.
3	Gained knowledge	12	Free-form descriptions of different knowledge related to the phishing and spoofing exercises.
4	Useful in working life	5	Descriptions mentioning the awareness of phishing and spoofing when working in an organization.
5	Inevitability	3	Descriptions of spoofing being commonplace, and how everyone should be aware of the topic.
6	Analysing message metadata	4	Descriptions of gaining hands-on skills for analysing message metadata to verify the authenticity of the sender.
7	Spoofing ease	13	Descriptions of how easy it is or how little skill it takes to create spoofed messages.

Figure 11: Thematic analysis table from co-authored article @mipro.hr [37]

There were a total of 39 responders, with 22 reports found to have statements relating to phishing and/or spoofing. The remaining reports were found to have content describing other areas of the course and have no bearing on this thesis.

Of the 22 reports examined in more detail, a total of 16 responses indicated raised awareness, 13 responses highlighted the ease of spoofing, and a further 12 responses showed that students had gained knowledge in this area which could be put to good use identifying fraudulent sites and emails.

### 5.3 Results of questionnaire

As stated earlier, only a small number of students responded to this voluntary feedback, twelve out of a potential twenty-six students. There was a total of six questions asked of the students, four of these questions required a mandatory response.

The results presented below are from the first four questions asked and are presented in excel formatting (Table 1-4), the full results as displayed in the original questionnaire are shown as supplementary information in appendix 3.

Table 1: Feedback question 1

Timestamp	Would you have preferred the option to do this lab at home using the Kotka environment?
03/01/2023 15:19:59	No
03/01/2023 15:23:13	Maybe
03/01/2023 15:38:51	Maybe
03/01/2023 15:40:20	No
03/01/2023 15:46:02	No
03/01/2023 15:46:49	No
03/01/2023 15:56:40	Maybe
03/01/2023 16:10:25	Yes
03/01/2023 18:49:18	No
03/01/2023 21:00:37	Maybe
03/01/2023 22:18:25	Yes
03/01/2023 22:28:31	No

Table 2: Feedback question 2

Timestamp	On a scale of 1 - 5 How successful was this lab in explaining the concept of phishing attacks and credential harvesting [1 - 5]
03/01/2023 15:19:59	3 - No opinion
03/01/2023 15:23:13	5 - Very
03/01/2023 15:38:51	4 - Somewhat
03/01/2023 15:40:20	4 - Somewhat
03/01/2023 15:46:02	4 - Somewhat
03/01/2023 15:46:49	5 - Very
03/01/2023 15:56:40	2 - A little
03/01/2023 16:10:25	4 - Somewhat
03/01/2023 18:49:18	4 - Somewhat
03/01/2023 21:00:37	4 - Somewhat
03/01/2023 22:18:25	5 - Very
03/01/2023 22:28:31	4 - Somewhat

Table 3: Feedback question 3

Timestamp	On a scale of 1 - 5 How successful was this lab in demonstrating browser warnings that could point to you being on a scam site? [1 - 5]
03/01/2023 15:19:59	4 - Somewhat
03/01/2023 15:23:13	5 - Very
03/01/2023 15:38:51	5 - Very
03/01/2023 15:40:20	5 - Very
03/01/2023 15:46:02	5 - Very
03/01/2023 15:46:49	5 - Very
03/01/2023 15:56:40	4 - Somewhat
03/01/2023 16:10:25	5 - Very
03/01/2023 18:49:18	5 - Very
03/01/2023 21:00:37	4 - Somewhat
03/01/2023 22:18:25	4 - Somewhat
03/01/2023 22:28:31	5 - Very

Table 4: Feedback question 4

Timestamp	On a scale of 1 - 5, did the lab make you re-examine your own online browsing behaviour? [1 - 5]
03/01/2023 15:19:59	1 - Not at all
03/01/2023 15:23:13	5 - A lot
03/01/2023 15:38:51	4 - Somewhat
03/01/2023 15:40:20	4 - Somewhat
03/01/2023 15:46:02	4 - Somewhat
03/01/2023 15:46:49	2 - A little
03/01/2023 15:56:40	4 - Somewhat
03/01/2023 16:10:25	5 - A lot
03/01/2023 18:49:18	5 - A lot
03/01/2023 21:00:37	1 - Not at all
03/01/2023 22:18:25	2 - A little
03/01/2023 22:28:31	4 - Somewhat

The following two questions were also asked of students to see if they could provide feedback that may help to improve the lab.

- 1) Suggestions on how the material or exercise could be improved.
- 2) Any other feedback they would like to give on the exercise.

No new information was found with these questions with each question only receiving a single response. Those responses can be found in Appendix 2, page 3.

## 6 DISCUSSION

From the thematic analysis the lab was indeed successful in raising awareness and informing students about how phishing attacks are performed. Upon examining the questionnaire though, students felt phishing and credential harvesting concepts were only “somewhat” explained.

As my first real attempt at creating an exercise from the ground up, it is perhaps too easy to look back in hindsight and identify what could have been changed or done differently. This exercise provides a foundation that can be built on in later iterations of this course.

Examining the lab instructions points to a lack of information or demonstration regarding credential harvesting. The previous version of the lab I had designed

had the students using a tool named blackeye in Linux. The software in question can automatically create a localhost version of multiple commonly used sites, e.g., Twitter, Facebook, Instagram, and Github to name a few. The problem is that we could be presenting students with too much information on how to perform phishing attacks, there is a very fine line between educating students to raise awareness and educating students on how to perform actual phishing attacks. At that point in time, it was also decided that demonstrating this software may increase the lab work beyond the allotted time, and the demonstration of such software may also fall on the wrong side of the educational line. A recent study has examined this very subject though, Hynninen [38 p.8] posits that "...even for novices in the field, there are seldom ethical issues regarding how to use hacking knowledge".

The exercise was performed relatively quickly by all students involved, this reveals that there is room to expand on the content of the lab, such content should probably include information, or demonstration of unencrypted transmissions over HTTP using the previously mentioned blackeye software.

There is always the option of attempting to gamify this area of education, as was discussed earlier in the thesis [6].

## **7 CONCLUSION**

Upon completion of the lab and having the opportunity to examine both the feedback from student written reports, and an anonymous feedback questionnaire, this work was found to be successful in raising awareness around the area of phishing.

Through a less conventional approach of educating people how these attacks are performed, we can engage with students at a fundamental level. This method, along with other labs done throughout the security fundamentals course, tasks students with making connections between the threats and how they could be used together to perpetrate attacks.



There is no single approach that can be taken that will suddenly eradicate phishing as an attack method. By continuing to evaluate how educators can communicate more effectively with students on the subject matter, this should reduce the risk of the student becoming a victim of this attack either in their personal or work life.

I believe that this is a field which would be detrimental to ignore and requires ongoing research and evaluation. The methods used by cyber criminals constantly evolves, and as such, the material and methods used need to be re-examined on a regular basis to ensure students are presented with relevant information, and that the level of engagement continues at the highest levels possible.

## REFERENCES

- [1] Phishing.org. History of Phishing. [No Date]. Phishing.org [Online] Available from: <https://www.phishing.org/history-of-phishing> [Accessed 24 February 2023]
- [2] Rapid7.com. Phishing Awareness Training. [No Date] Rapid7.com [Online]. Available from: <https://www.rapid7.com/solutions/phishing-awareness-training/> [Accessed 24 January 2023]
- [3] mimecast.com. Phishing Awareness Training. [No Date] mimecast.com. [Online]. Available from: <https://www.mimecast.com/content/phishing-training/> [Accessed 24 January 2023]
- [4] sans.org. Robust Phishing Awareness Simulation Training that Changes Behavior. [No Date]. sans.org. [Online] Available from: <https://www.sans.org/security-awareness-training/products/security-awareness-solutions/phishing/> [Accessed 24 January 2023]
- [5] M. Bada et al. Cyber Security Awareness Campaigns: Why do they fail to change behaviour? 09 January 2019. Arxiv [PDF] Available from: <http://arxiv.org/abs/1901.02672> [Accessed 25 January 2023]
- [6] Z. A. Wen et al. What.Hack: Engaging Anti-Phishing Training Through a Role-playing Phishing Simulation Game. CHI '19: CHI Conference on Human Factors in Computing Systems. Glasgow, Scotland, UK. ACM. p.2. 2019. [PDF] Available from: [https://www.cs.cornell.edu/~eland/papers/chi2019\\_whathack.pdf](https://www.cs.cornell.edu/~eland/papers/chi2019_whathack.pdf) [Accessed 27 March 2023]
- [7] D. Winder. Cisco Hacked: Ransomware Gang Claims It Has 2.8GB Of Data. 13 August 2022. Forbes. [Online] Available from: <https://www.forbes.com/sites/daveywinder/2022/08/13/cisco-hacked-ransomware-gang-claims-it-has-28gb-of-data/?sh=699977644043> [Accessed 9 February 2023]
- [8] E. Shilling. The 9 Lives of the Spanish Prisoner, the Treasure-Dangling Scam That Won't Die. 03 August 2016. Atlas Obscura. [Online] Available from: <https://www.atlasobscura.com/articles/the-9-lives-of-the-spanish-prisoner-the-treasure-dangling-scam-that-wont-die> [Accessed 28 March 2023]
- [9] APWG. Phishing Activity Trends Report, 1<sup>st</sup> Quarter 2012. 19 July 2012. APWG. [PDF] Available from: [https://docs.apwg.org/reports/apwg\\_trends\\_report\\_q1\\_2012.pdf](https://docs.apwg.org/reports/apwg_trends_report_q1_2012.pdf) [Accessed 29th March 2023]
- [10] APWG. Phishing Activity Trends Report, 3<sup>rd</sup> Quarter 2022. 12 December 2022. APWG. [PDF] Available from:

[https://docs.apwg.org/reports/apwg\\_trends\\_report\\_q3\\_2022.pdf](https://docs.apwg.org/reports/apwg_trends_report_q3_2022.pdf) [Accessed 29 March 2023]

[11] M. Leonhardt. 'Nigerian prince' email scams still rake in over \$700,000 a year—here's how to protect yourself. 18 April 2019. <http://www.cnbc.com>/CNBC LLC. [Online] Available from: <https://www.cnbc.com/2019/04/18/nigerian-prince-scams-still-rake-in-over-700000-dollars-a-year.html> [Accessed 31 March 2023]

[12] R. Alli et al. Detecting advance fee fraud emails using self-referential pronouns: A preliminary analysis. 2018. University of Portsmouth. [PDF] Available from: [CRAIG 2018 cright AF Detecting advance fee fraud emails using self referential pronouns.pdf \(port.ac.uk\)](CRAIG_2018_cright_AF_Detecting_advance_fee_fraud_emails_using_self_referential_pronouns.pdf(port.ac.uk)) [Accessed 31 March 2023]

[13] E. Williams et al. Exploring susceptibility to phishing in the workplace. International Journal of Human-Computer Studies. 2018 [Online] Available from: <https://www.sciencedirect.com/science/article/pii/S1071581918303628> [Accessed 31 March 2023]

[14] A. Popa. Spear phishing: top 6 biggest attacks in history. 11 August 2021. Attack Simulator. [Online] Available from: <https://attacksimulator.com/blog/biggest-spear-phishing-attacks> [Accessed 04 April 2023]

[15] S. Nasralla. Austria's FACC, hit by cyber fraud, fires CEO. 25 May 2016. Reuters. [Online] Available from: <https://www.reuters.com/article/us-facc-ceo-idUSKCN0YG0ZF> [Accessed 17 April 2023]

[16] Traficom. Another wave of FluBot: malware being spread by SMS. 10 May 2022. Traficom. [Online] Available from: [https://www.kyberturvallisuuskeskus.fi/en/varoitus\\_1/2022](https://www.kyberturvallisuuskeskus.fi/en/varoitus_1/2022) [Accessed 17 April 2023]

[17] S. Sjouwerman. Phishing Attack in Finland Uncovers Sophisticated Smishing Scheme. 20 July 2022. KnowBe4, Inc. [Online] Available from: <https://blog.knowbe4.com/phishing-attack-in-finland-uncovers-sophisticated-smishing-scheme> [Accessed 17 April 2023]

[18] Vero. Scam messages. 04 April 2023. Vero. [Online] Available from: <https://www.vero.fi/en/About-us/contact-us/efil/Information-on-e-services/scam-messages> [Accessed 17 April 2023]

[19] L. Irwin. What Is Angler Phishing? Definition, Examples & Prevention. 24 June 2019. IT Governance UK blog. [Online] Available from: <https://www.itgovernance.co.uk/blog/beware-of-angler-phishing> [Accessed 25 April 2023]

- [20] APWG. Phishing Activity Trends Report, 1<sup>st</sup> Quarter 2021. 08 June 2021. APWG. [PDF] Available from: [https://docs.apwg.org/reports/apwg\\_trends\\_report\\_q1\\_2021.pdf](https://docs.apwg.org/reports/apwg_trends_report_q1_2021.pdf) [Accessed 25 April 2023]
- [21] Valimail. Phishing vs Pharming. [No Date] Valimail. [Online] Available from: <https://www.valimail.com/guide-to-phishing/phishing-vs-pharming> [Accessed 24 April 2023]
- [22] <https://www.microsoft.com/en-us/security/business/security-101/what-is-business-email-compromise-bec>
- [23] TechTarget. Definition pharming. April 2021. TechTarget. [Online] Available from: <https://www.techtarget.com/searchsecurity/definition/pharming> [Accessed 08 May 2023]
- [24] Kaspersky. What Is Pharming and How to Protect Yourself. 13 September 2017. Kaspersky [Online] Available from: <https://www.kaspersky.com/resource-center/definitions/pharming> [Accessed 05 May 2023]
- [25] Various Contributors. Hyper-V architecture. Updated 03 May 2023. Learn.microsoft.com [Online] Available from: <https://learn.microsoft.com/en-us/windows-server/administration/performance-tuning/role/hyper-v-server/architecture> [Accessed 05 May 2023]
- [26] XCP-ng.org. XCP-ng documentation. Updated 15 June 2022. XCP-ng.org [Online] Available from: <https://xcp-ng.org/docs/> [Accessed 05 May 2023]
- [27] VMware.com. VMware ESXi. [No Date]. VMware.com [Online] Available from: <https://www.vmware.com/nordics/products/esxi-and-esx.html> [Accessed 05 May 2023]
- [28] VMware.com. [No Date]. VMware Workstation Pro. [Online] Available from: <https://www.vmware.com/nordics/products/workstation-pro.html> [Accessed 05 May 2023]
- [29] VirtualBox.org VirtualBox. [No Date].VirtualBox.org [Online] Available from: <https://www.virtualbox.org> [Accessed 05 May 2023]
- [30] M. Siegrist. Leveraging Virtualization Technology for e-Learning. 10 January 2011. CloserStill Media, Ltd. [Online] Available from: <https://www.LearningGuild.com/articles/615/leveraging-virtualization-technology-for-e-learning> [Accessed 08 May 2023]
- [31] K. Pyykkö. Using virtual machines to create standardized learning environment for software engineering

- courses. MSc Thesis. Lappeenranta-Lahti University of Technology LUT. 2022. [PDF] Available from: <https://lutpub.lut.fi/bitstream/handle/10024/164801/Using%20virtual%20machines%20to%20create%20standardized%20learning%20environment%20for%20software%20engineering%20courses%20-%20Karri%20Pykk%C3%B6.pdf?sequence=1> [Accessed 08 May 2023]
- [32] M. Ashfaq. Virtual machines In Education. Master's Thesis. Oslo University College. 2007. [PDF] Available from: <https://www.duo.uio.no/handle/10852/9701> [Accessed 08 May 2023]
- [33] Ubuntu.com. Install Ubuntu Desktop. [No Date] Ubuntu.com. [Online] Available from: <https://ubuntu.com/tutorials/install-ubuntu-desktop#1-overview> [Accessed 08 May 2023]
- [34] A. Garnett et al. Apache vs Nginx: Practical Considerations. 08 January 2015. DigitalOcean. [Online] Available from: <https://www.digitalocean.com/community/tutorials/apache-vs-nginx-practical-considerations> [Accessed 09 May 2023]
- [35] A. Garnett. How To Install Nginx on Ubuntu 22.04. 26 April 2022. DigitalOcean. [Online] Available from: <https://www.digitalocean.com/community/tutorials/how-to-install-nginx-on-ubuntu-22-04> [Accessed 09 May 2023]
- [36] Entrust. What is a self-signed certificate? [No Date] Entrust Corporation. [Online] Available from: <https://www.entrust.com/resources/faq/what-is-a-self-signed-certificate> [Accessed May 10 2023]
- [37] A. Kerr et al. (2023). "Towards Improving Online Security Awareness Skills with Phishing and Spoofing Labs." [Manuscript accepted for publication] In the Proceedings of the Mipro 2023 ICT and Electronics Convention, MIPRO 2023 Hybrid Convention [Accessed 10 May 2023]
- [38] T. Hynninen. (2023). "Student perceptions of ethics in cybersecurity education." [Manuscript accepted for publication] In the Proceedings of the 2023 Conference on Technology Ethics, Tethics 2023. [Accessed 17 May 2023]

IT00CO77 Security Fundamentals / OT00EK08 Ohjelmistojen tietoturva  
You may complete the exercises in groups of 1-3 people. All group members should contribute equally

### Week 9 exercise: A phishing expedition

#### Phishing for beginners!



#### Introduction and preparation

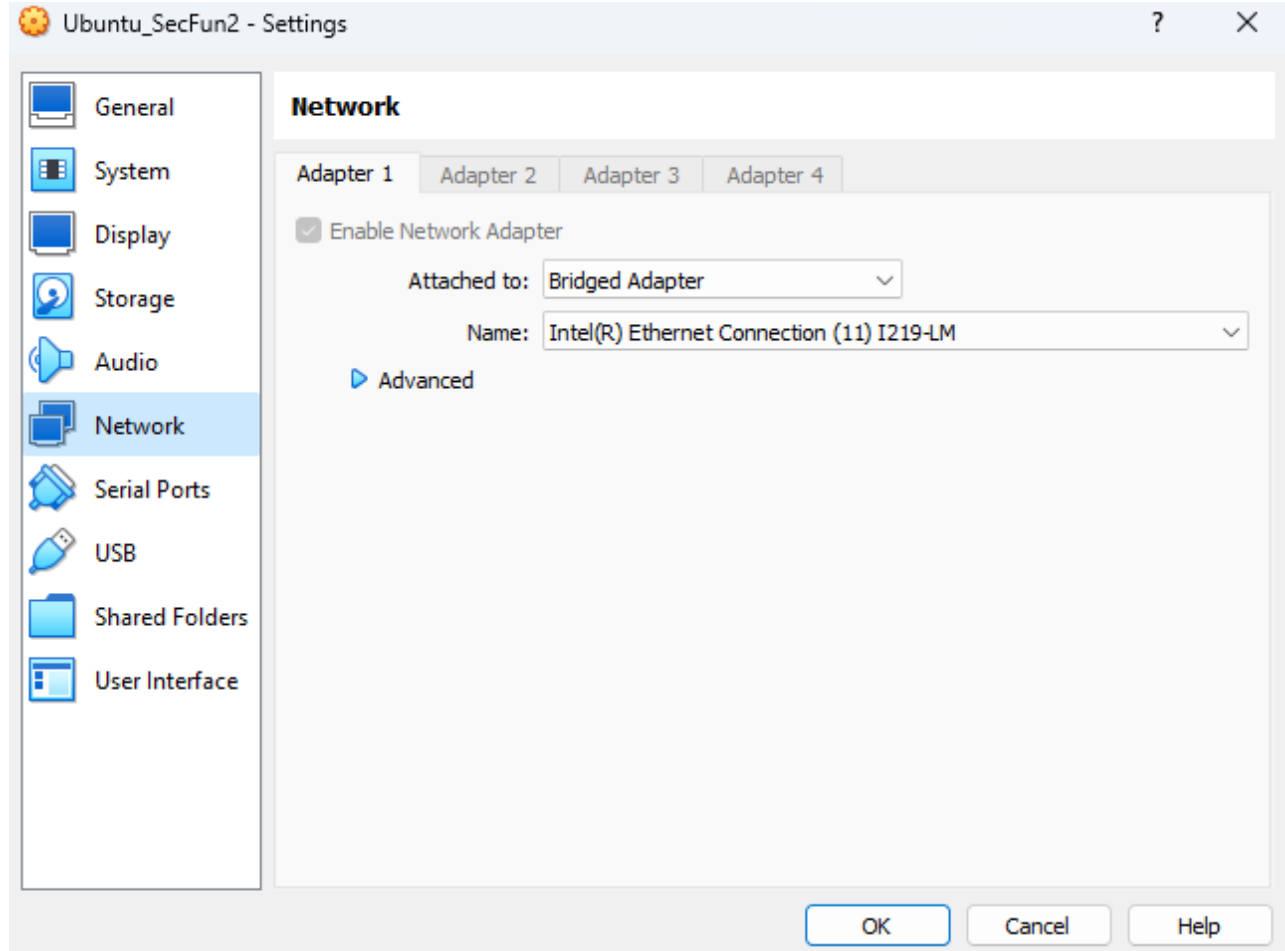
**These exercises will be performed using the classroom PCs**, they can be done in groups of 1-3. If you are not on campus you may try the first part of it in the Virtual Lab as well but we currently we have no proper support for this. Remember that the whole group should answer the questions personally to ensure that everyone understands the concepts being discussed.

If any of the commands fail to work, try adding sudo in front of them (do this with caution, though. Not



everything is meant to be done with administrator privileges).

Open VirtualBox and import the latest version of the Security fundamentals VM from “\Public\Security Fundamentals VM folder”. Before you start the VM, click on settings, go to the network tab, in adapter 1 Name: Click on the dropdown menu and select the first available adapter, this will vary depending on your location.



Once the VM reaches the login screen, select user “cisco” and the password is just “password”.

### Browsing the internet

Inside the VM, open the terminal and issue the command “ip address” and take a note of your IP. It should be in the format: 172.20.x.x since we have a bridged connection, the VM will have its own IP.

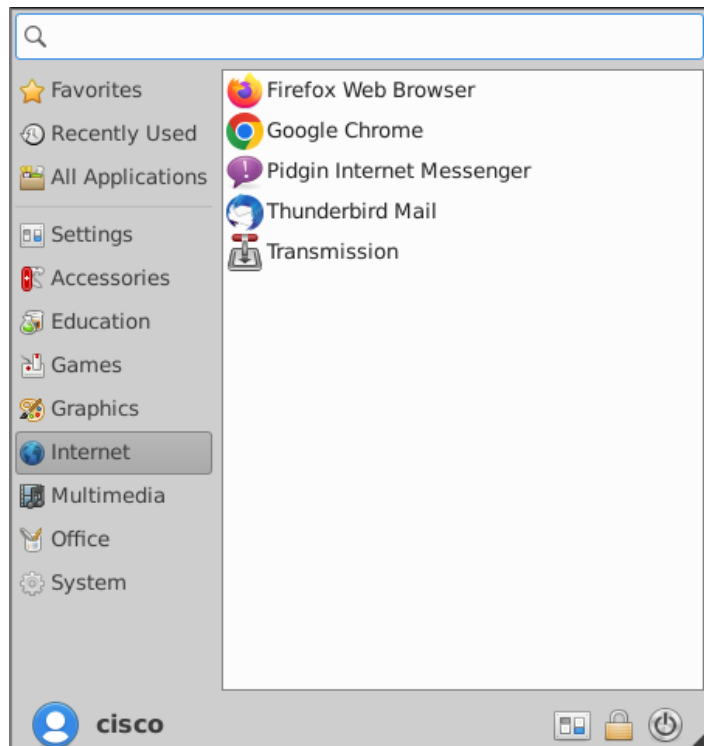
### Open up google chrome

Try navigating to some sites like [www.youtube.com](http://www.youtube.com) or [facebook.com](http://facebook.com) to ensure that chrome and the internet are working correctly, if you browse to YouTube, you should be able to play video with sound in MB316.

Navigate to [web2.com](http://web2.com), a web form should be displayed. **This is currently hosted on a web server we have running on campus.**

Type [learn.xamk.fi](http://learn.xamk.fi) to the address bar and try logging in. If there are any problems, be sure to let either Alex/Jevgeni or Timo know.

Minimize google chrome and from the top left menu select the internet tab and choose Firefox this time.





Now navigate back to learn by typing learn.xamk.fi in the address bar.

Open chrome back up and navigate back to learn.xamk.fi.

**Are there any differences in what information the browser is presenting to you? If so, what are the differences?**

Try typing https://learn.xamk.fi into the address bar for both browsers. Click on the advance section and read the information provided.

**What is the warning provided by the browsers?**

**What does the warning actually mean?**

Use the classroom pc web browser (outside the VM) and type learn.xamk.fi into the address bar. Does the classroom PC connect to the correct site? Obviously there is an issue in our VM somewhere.

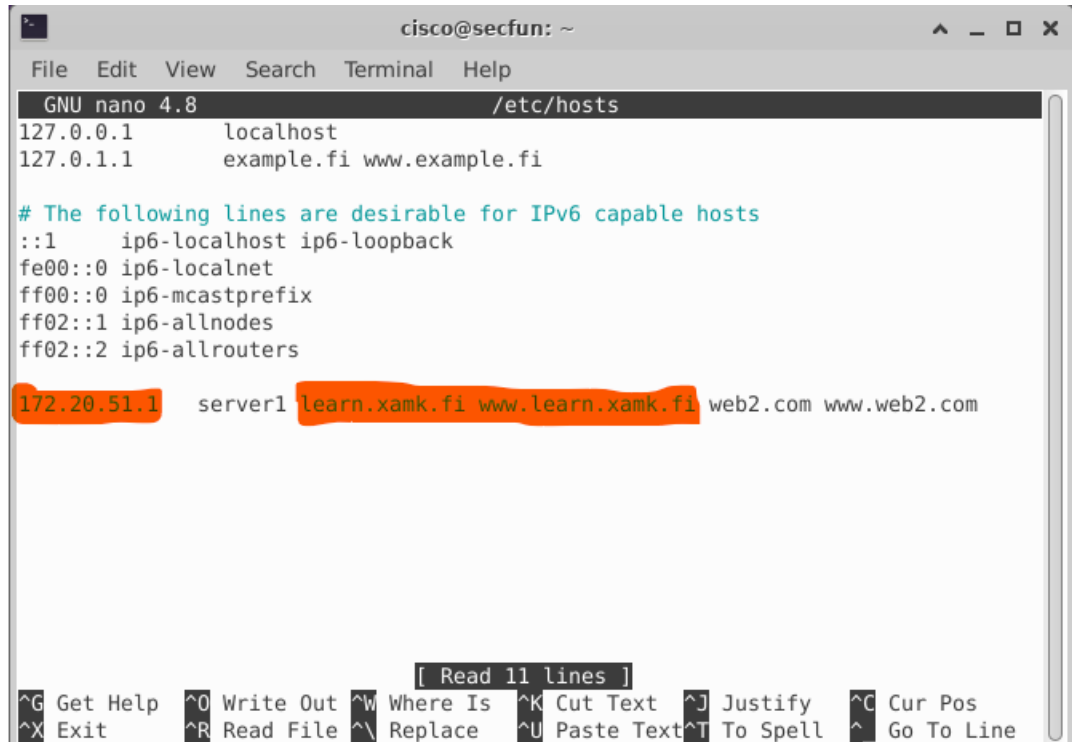
**Google for “what is the hosts file used for in Linux”, what service is being provided by the hosts file?**

**From the information you have so far gathered, what is happening when we try to navigate to learn.xamk.fi in the Linux VM?**

With a little better understanding of what the hosts file is, we can now remedy the current situation that the VM is redirecting traffic destined for learn.xamk.fi.

Open terminal, type the following command:  
sudo nano /etc/hosts

Arrow down to the line that includes learn.xamk.fi and www.learn.xamk.fi. Delete only those 2 references from the line, press ctrl and x to exit, press y to save the modified buffer, then press enter to save the file with its current name.



```
cisco@secfun: ~
File Edit View Search Terminal Help
GNU nano 4.8 /etc/hosts
127.0.0.1 localhost
127.0.1.1 example.fi www.example.fi

# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters

172.20.51.1 server1 learn.xamk.fi www.learn.xamk.fi web2.com www.web2.com

[ Read 11 lines ]
^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos
^X Exit ^R Read File ^\ Replace ^U Paste Text ^T To Spell ^_ Go To Line
```

Once you have removed the offending text from the hosts file, open up Chrome or Firefox and try navigating back to learn.xamk.fi, the correct page should now load.

### A targeted phishing expedition

As was just briefly mentioned, it is possible to quickly set up a credential harvesting site in minutes, what if we wanted to target something more specific instead of a social media platform? Let's quickly throw up our own website in our VM, this just demonstrates the site creation process, it will not cover how to harvest credentials.

Inside the VM, open a web browser and type example.fi into the address bar.

The page is being hosted on your VM using Nginx. Nginx functions as a load balancer/web server/reverse proxy. The page you just visited in chrome is located at

/var/www/example.fi/html/index.html

The first step would be to target a specific site.... perhaps peppi.xamk.fi for example.

Some pages (like linkedin.com) can be downloaded directly using the wget command from terminal, Peppi is not one of these pages.

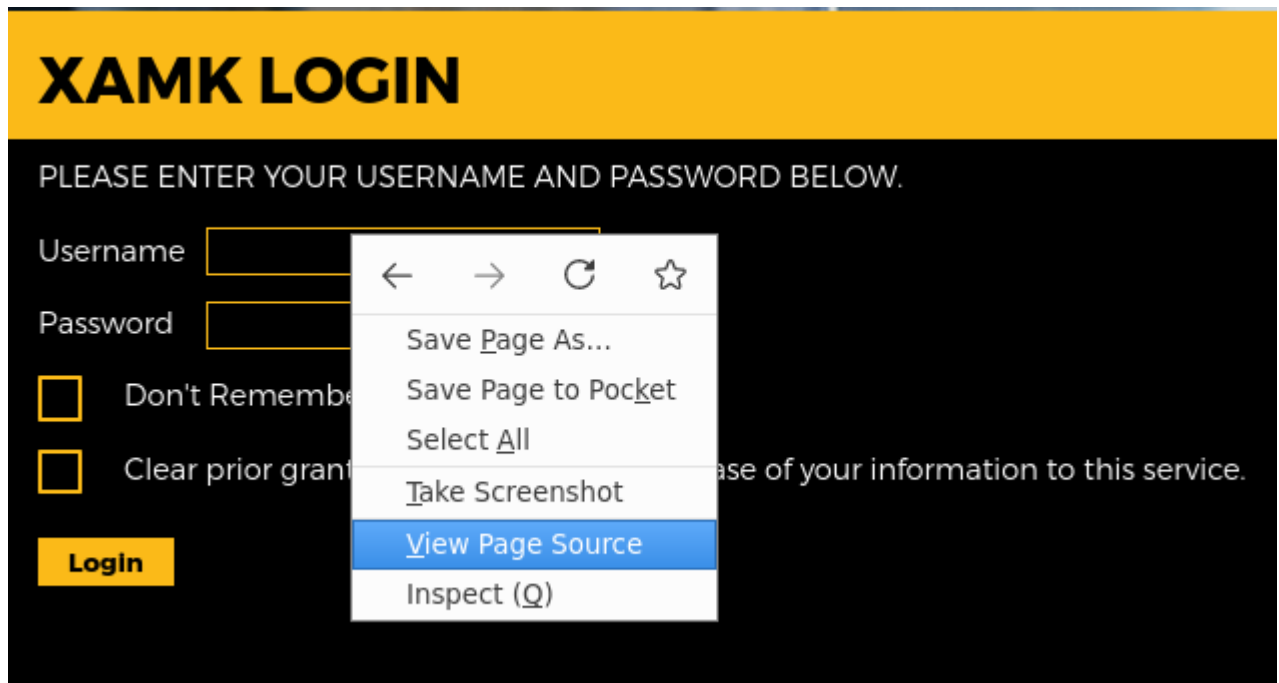
Take a look at this link

<https://blog.hubspot.com/marketing/web-design-html-css-javascript>

**Can you give a brief definition of the following elements of a website? HTML/JavaScript/CSS**

### **Phishing for beginners**

First double click on mousepad to open a blank text file, then open a web browser in the VM and navigate to peppi.xamk.fi. Right click somewhere in the black space on the page and select view source.



In the new tab that has opened, press ctrl + a, select all, then press ctrl + c to copy it.

Go back to your open blank text file and press ctrl + v to paste the html code. Save the file as index.html in the default location and close the file.


Open terminal and navigate to /var/www/example.fi/html/ and type the following command:

```
cp ~/index.html  
/var/www/example.fi/html/index.html
```

This will copy your html file to the root directory of example.fi.

Try opening a web browser and go to example.fi now.

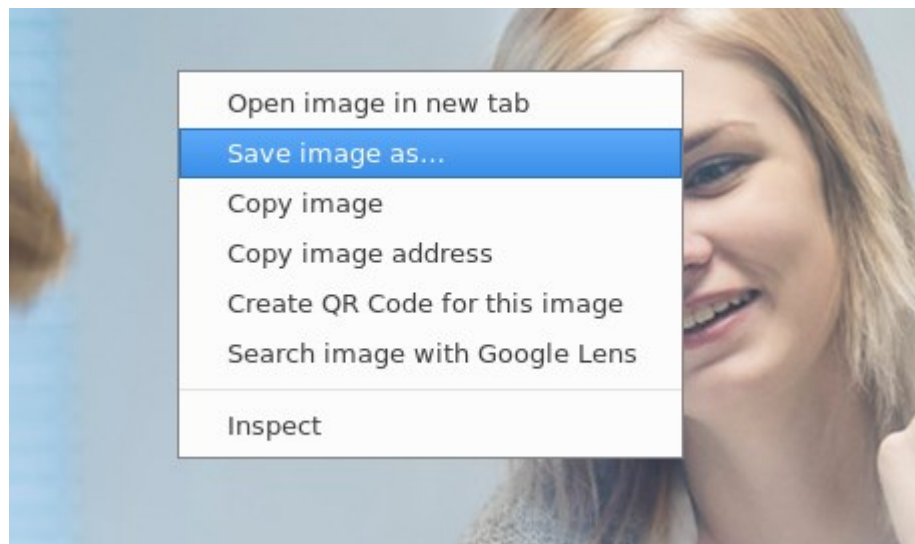
---

 Xamk logo  
XAMK LOGIN  
PLEASE ENTER YOUR USERNAME AND PASSWORD BELOW.  
Username   
Password   
 Don't Remember Login  
 Clear prior granting of permission for release of your information to this service.

Not quite what we are looking for. Right now we only have HTML, We have no images or styles linked in to our web page. We need at least 2 more things to get something resembling the peppi login page. The first thing we need is the Xamk logo image, the second thing we need is the css styling.

Open up mousepad again.

Open peppi.xamk.fi in the web browser, right click the image above the yellow XAMK LOGIN bar and then select save image as...



Rename the file logo.jpg and save it in the default location.

View page source again (right click in black space).  
Near the top of the code, we can see a link to the css  
file

```
<!DOCTYPE html>
<html>
<head>
<meta charset="utf-8">
<meta name="viewport" content="width=device-width, initial-scale=1, maximum-scale=1, user-scalable=no">
<title>Web Login Service</title>
<link rel="stylesheet" type="text/css" href="/idp/css/xamk-2022_1.css">
</head>
```

Left click on the link to the css file. Again in the new  
tab, press ctrl + a then ctrl + c. Paste the code to our  
empty mousepad file and save the file as  
example.css in the default location.

Open terminal and navigate to  
/var/www/example.fi/html  
See what is inside this directory.

```
cisco@secfun:/var/www/example.fi/html$ ls
css  images  index.html
```

We are now going to copy the logo.jpg file to the  
images folder and then copy the example.css to the  
css folder with

```
cp ~/Downloads/logo.jpg ./images/logo.jpg
&& cp ~/example.css ./css/example.css
```

From the /var/www/example.fi/html folder, change  
directory to images and type ls to ensure we have  
the logo.jpg in the correct folder.

```
cd images
ls
```

Then navigate back to the html folder and again  
change directory to the css and type ls to ensure we  
have the example.css in the current folder.

```
cd .. && cd css
ls
```

You should now have both these files in the proper place. Now we just need to link them to our index.html page.

Use the following command to change directory to /var/www/example.fi/html and edit the index.html page:

```
cd .. && sudo nano index.html
```

```

cisco@secfun: /var/www/example.fi/html/images
GNU nano 4.8 index.html
!DOCTYPE html>
<!-- saved from url=(0066)https://xidp.xamk.fi/idp/profile/SAML2/Redirect/SSO?execution=els2 -->
<html><head><meta http-equiv="Content-Type" content="text/html; charset=UTF-8">

<meta name="viewport" content="width=device-width, initial-scale=1, maximum-scale=1, user-scalable=no">
<title>Web Login Service</title>
<link rel="stylesheet" type="text/css" href="./css/example.css">
</head>
<body>
<header>
  <div id="top">
    
  </div>
</header>

```

Edit the highlighted sections so they match the above text. Once you are done, press ctrl + x to exit, press y to save changes, and press enter to overwrite the index.html file.

If you now browse to example.fi we should see a decent copy of the page in place. If we managed to get traffic to visit our site, we could add some code that would log any credentials typed to a file.

**What line would we add to the targets hosts file to redirect them to our fake website?**

**Currently we can only connect via http, what would we need to enable https connection?**

At the end, consider the following:

1. There are lots of applications available which will quickly build common social media login pages in an

attempt to harvest credentials. Think about how many of these you are registered to, examples like YouTube, Facebook, LinkedIn, Twitter and so on. How many of these use your email as the user, do

you have the same password for more than one of these sites?

2. If a malicious actor was to get a combination of your email and password for one site, would they then be able to login to multiple other places?

### **Optional: Only HTTP?**

We can only connect to our site via http. If you attempt to connect via https the connection will automatically be refused. If you have finished the previous exercises fairly quickly, you can attempt this part so that we can enable an https connection but is entirely optional.

Ensure nginx openssl is installed:

```
sudo apt update && sudo apt install nginx openssl
```

Create the required directory:

```
sudo mkdir /etc/nginx/certificates
```

Change to the new directory:

```
cd /etc/nginx/certificates
```

Create the self-signed web certificate and the private key:

```
openssl req -new -newkey rsa:4096 -x509 -sha256 -days 365 -nodes -out nginx-certificate.crt -keyout nginx.key
```

The only important field here is the “Common Name”, you should assign the IP address of your VM to this field.



We need to link our certificate/key to our site by editing our sites-available configuration file.

```
sudo nano /etc/nginx/sites-available/example.fi
```

Edit the file so it resembles the screenshot below:

```
server {
    listen 80;
    listen [::]:80;

    listen 443 ssl;
    listen [::]:443 ssl;
    ssl_certificate /etc/nginx/certificates/nginx-certificate.crt;
    ssl_certificate_key /etc/nginx/certificates/nginx.key;

    root /var/www/example.fi/html;
    index index.php index.html index.htm index.nginx-debian.html;
```

Save and exit, test the syntax with “sudo nginx -t” and then restart nginx:

```
sudo systemctl restart nginx
```

Test your browser is accessing the site by typing “example.fi” into the browser and then test “https://example.fi”. This should show a security issue as the certificate is self-signed.

These warnings can be suppressed, and if we have access to alter the targets hosts files, we should also be able to suppress these warnings in their browsers.

Would you have preferred the option to do this lab at home using the Kotka environment?

12 responses

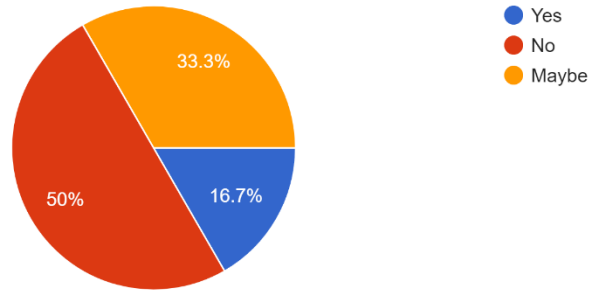


Figure 12: Question 1 for feedback

On a scale of 1 - 5 How successful was this lab in explaining the concept of phishing attacks and credential harvesting

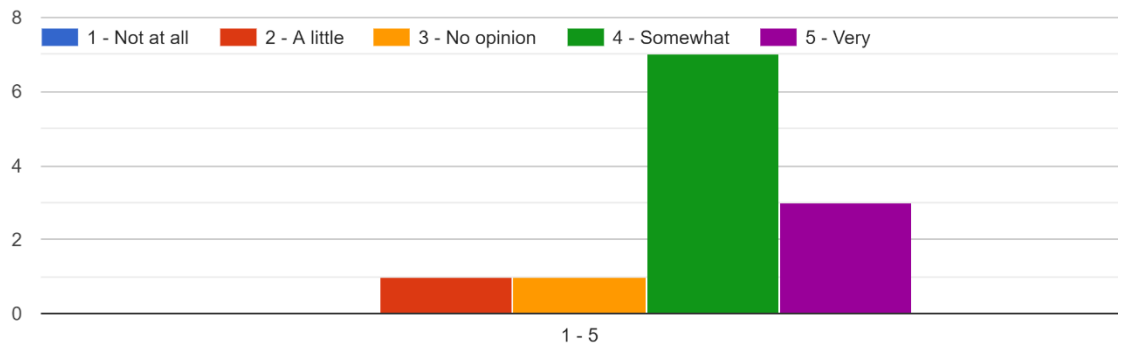


Figure 13: Question 2 for feedback

On a scale of 1 - 5 How successful was this lab in demonstrating browser warnings that could point to you being on a scam site?

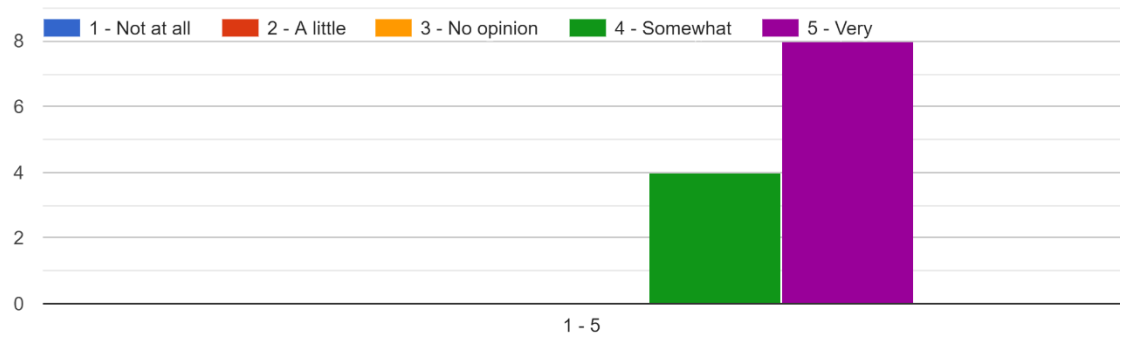


Figure 14: Question 3 for feedback

On a scale of 1 - 5, did the lab make you re-examine your own online browsing behaviour?

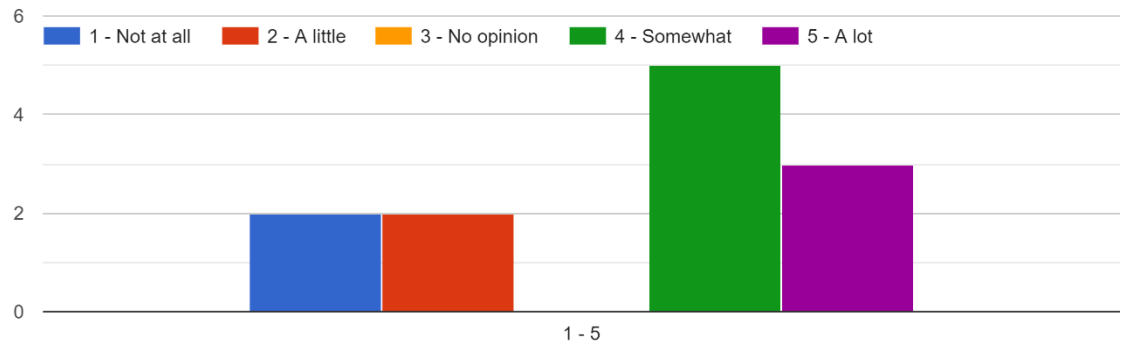


Figure 15: Question 4 for feedback

Do you have suggestions how this lab be improved in regards to information presented, or the practical exercise?

1 response

It would be convenient to have an option to finish task at home

---

Is there any other feedback you would like to give regarding the lab?

1 response

The lab explained the danger quite fully and understanding.

Figure 16: 2 general questions for feedback