5 YEARS AGAINST PHISHERS: ATTACK TECHNIQUES, PROTECTION DESIGN AND LESSONS LEARNED

Third ESCAPE Seminar report

FIRRERA Manuel s329226

6 December 2023

1 Introduction

This report covers all the topics presented by Stefano Traverso, a cyber-security expert and co-founder of the ”Ermes Security” company, for the third seminar of the ESCAPE cycle.
The talk focused on giving more insights on the **phishing** cyber-scam, from the techniques used by attackers to detection strategies.
Additionally, Mr. Traverso showcased his company’s efforts in developing malicious website detection systems employing artificial intelligence algorithms, which represent the latest frontier in the field of computing.

2 Social engineering

Before talking about phishing threats it is mandatory to describe a specific technique known as **social engineering**, which is usually leveraged for this kind of attacks.
Social engineering is defined as a set of actions used to deceive individuals into leaking normally confidential information, such as keys or credentials.
It can also be used to induce victims into performing actions or influencing their decisions.
Social engineering is usually performed where other types of attacks are not possible or time consuming.
Citing the host, it can be subdivided in the following steps:

- **Prepare**: gathering background information about the victims.
- **Infiltrate**: establishing a relationship with the victim and building trust.
- **Exploit**: leveraging trust and weaknesses to induce the victim to perform actions.
- **Disengage**: gracefully ending involvement, possibly leaving no traces behind.

Social engineering leverages human traits such as:

- **Emotions**: fear, excitement, curiosity, anger, guilt or sadness, used to convince victims.
- **Urgency**: Time-sensitive opportunities, request for immediate attention, to induce victims into error.
- **Trust**: confidence and believability to avoid rising any suspicions.

As demonstrated in the ’Cost of a Data Breach Report 2023’ by the Ponemon Institute and IBM, such attacks incurred recovery costs and reputational damage amounting to approximately 4 million USD that year, surpassing every other means of attack.
For this reason Social engineering should be considered as a real threat to company security and should be addressed with maximum priority.
As already said, this report will focus solely on phishing, which is one of the many social engineering type of attacks.
In the same category we can find baiting, whaling, quid pro quo to name a few.

3 Phishing

Phishing can be described by the threat model in figure 1, taken from Maltego website, a cyber-investigative company.

Through the usage of a **phishing kit**, an attacker crafts a fake website or attachment which is then distributed via a phishing campaign, which can involve e-mails, social media posts and other means of digital communication.
The possible outcomes are numerous, starting from malicious configurations to financial frauds, business theft and/or data exfiltration.

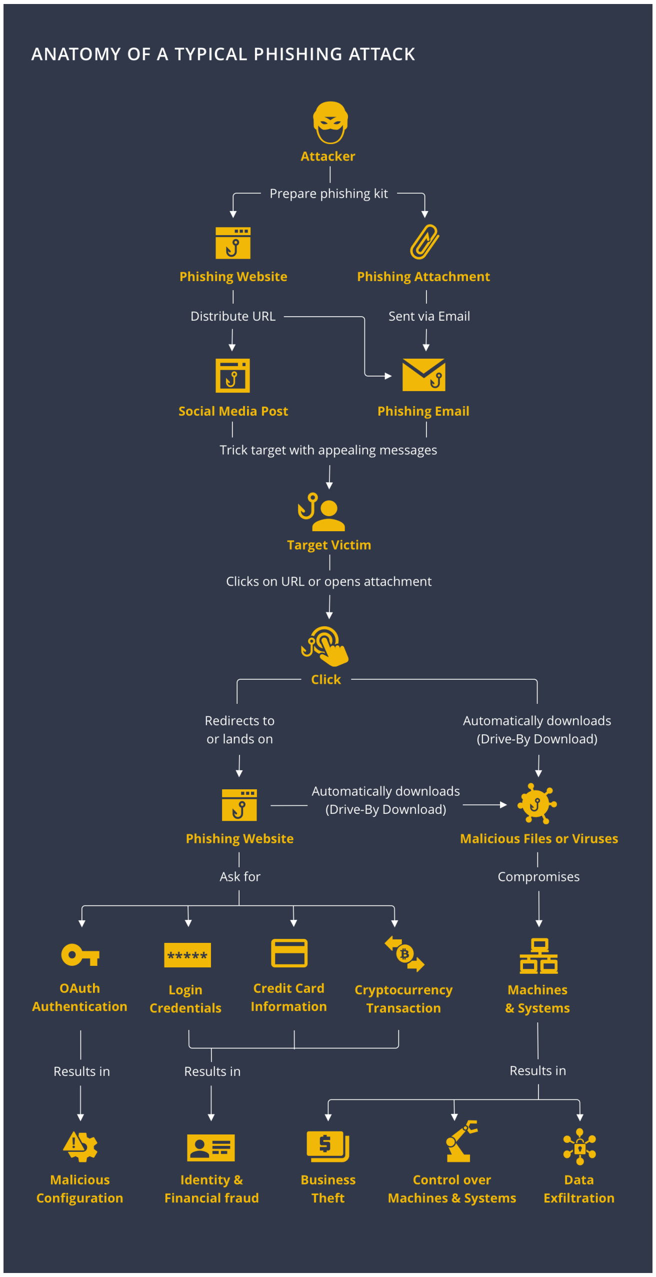

Figure 1: Phishing threat model

3.1 Creating a phishing campaign

In the first phases of the attack, hackers typically conduct recoinnaissance operations using tools such as *hunter.io* and *crosslinked* to get mailing lists, mailing conventions and other valuable pieces of information. These tools are provided with databases filled by continuously scanning the web to get the maximum amount of information possible.
Then the campaign can be created with toolkits such as *GoPhish* and *Brevo*
A website front-end is then built using *zPhisher* or *SocialFish*, as websites front-end is always publicly available, making this process extremely simple.
Once these components are ready, a domain name must be chosen. The objective is to trick the victim into believing the authenticity of the newly crafted site, so the URL domain name has to be similar, by looks (FaceboOk.com), by subdomain name (facebook.fakesite.com) or even by assonance (feisbuc, italian pronunciation); this is defined as **squatting**.
Lastly an HTTPS certificate can be provided. This can be done for free thanks to a number of "free Certification Authorities" as *Let's encrypt*. At this point the campaign can start, mails containing the malicious URL are finally sent to the victims and the attackers must simply wait for their victims to do the rest.

3.2 Protecting the crafted website

The web is continuously indexed by automatic programs called **spiders** or **crawlers**, for purposes like search engines data collection or security verification. To avoid these checks, malicious actors implement a series of countermeasures:

- **Anti-indexing**: crawlers can usually be stopped by declaring the content as not-indexable in a specific configuration file on the website.
- **Cloaking**: The website performs checks before processing the incoming request in order to avoid attracting unwanted attention.
  Cloaking can be further divided into:
  - **Server-side cloaking**: blocks navigation by looking at data contained in traffic carrying HTTP requests such as User Agent and IP.
  - **Client-side cloaking**: blocks navigation by looking at browser data or user actions, including mouse movement, clicks and cookies.

  - **Hybrid cloaking**: combines both techniques.

4 Advanced-phishing

More advanced techniques can be employed to enhance the effectiveness of a phishing campaign.
These methods usually involve complex procedures and require deep knowledge of the softwares/protocols used by the victim.
Mr. Traverso described the following three attacks:

- **HTML smuggling**: it's a technique based on inserting some malicious payload within an HTML attachment, that is sent for example via e-mail. This is a very good way to fool the user to open it and to get the malware.

- **Evil Proxy**: this technique is used to bypass multifactor authentication by capturing MFA access tokens and it's based on man-in-the-middle attacks (MITM).

- **Browser in the browser**: it's a technique that exploites pop-up windows for SSO authentication, by replicating a login window living in the context of the webpage, where the user can be easily fooled by inserting credentials and passwords.

These are some among the many possible attacks possible, which can trick even the most trained individuals.

5 Detection

Phishing websites can usually be detected by performing various checks on it's structure and components.

- **URL**: Squatting patterns, different domain name from the brand's one, unusual TLDs or ports can give away the true website nature.

- **Certificates**: Free certifications authorities such as, *let's encrypt*, *zerossl*, *sslforfree* are never used by legit websites as they do not guarantee their authenticity.

- **Background information**: To get more information about the domain name, databases such as the *WHOIS* one can be used. The database will give information about the registrar, registrant, organization country and much more, although recently, due to privacy regulations, less info are set to public use.

- **Source code analysis**: The suspicious website can be further analyzed to search for external links, broken elements, abnormal console logs which usually are a dead giveaway of a phishing website. *Wappalyzer* can be used for this purpose.

- **User interface check**: A phishing website page is usually badly built, as it does not need to stand any regulation authority approval.
  Missing cookie banners, suspicious cookies, JS variables, missing privacy policies and security headers are something to look for.

- **DNS Analysis**: By using powerful tools such as *dnsdumpster.com*, we can get the domain IP address, geolocation and DNS mapping. Any suspicious entry, location or mapping can be used to alert the user.

Thankfully numerous automatic checking tools already exist to automate the detection process.
VirusTotal is a tool used to detect malwares by scanning the files with 70 different anti-virus engines.
URLscan.io, a website scanning tool, is another example.
These services deeply scan the target by using the techniques mentioned above.

5.1 Phishing kit detection

Security professionals can detect/search specifically for phishing kits using opensource tools such as *kitphishr* or *stalkphish* to understand how to attackers craft their websites, get mailing lists, identify people.
They can then act consequently by employing specific countermeasures, instructing all the involved personnel.
Authors of the kits can sometimes be found by reverse engineering the source codes. They usually include pieces of code into the kit to try and get data as well as the attackers.

5.2 AI-Based detection

In conlusion, Mr. Daniele Traverso showed it's company researches in the artificial intelligence field as it can play an important role in the phishing detection

scenario.
An intelligent algorithm can check on a website features to try and discern between benign/malign nature.
A good crawler or spider is needed to bypass cloaking, anti-indexing and other countermeasures to finally let the AI do it's job.

- **PROs**: scalability, as the internet is always growing, 10-30M new domains per day.
- **CONs**: misclassifications, which can cause denial of service for the final user. This aspect was crucial for Ermes as slightly unprecise classification can lead to thousands of wrongly marked websites.

6 Conclusions & personal considerations

In conclusion we saw that phishing is a really simple attack yet one of the most effective as it leverages the human factor, which is usually the weakest link in the chain of cyber-security.
This is because phishing leverages human psychology through social engineering tactics.
Thanks to tools and open source code it has become highly ”democratized”, making it possible for individuals with minimal technical expertise to launch sophisticated phishing campaigns.
This accessibility means more significant financial and credibility losses for the companies.
For these reasons organizations ranging from small businesses to big corporates should heavily invest into security design and auditing as well as educating all personnel into following a list best practices.
For example, it could be useful to conduct internal phishing attacks, to monitor how the employees behave in case of attacks. This practices are known as **red teaming operations**.

Personally I found the talk very interesting as it covered a broad range of subjects, going into the right amount of details.
It highlighted the fact that, as security professionals, we must raise awareness about the threats of the interconnected world we live in.
It also emphasized the double-sided nature of the now ubiquitous AI.
While it can be greatly helpful in recognizing malicious code and websites, it can also give an advantage to the attackers, as AI-generated content is still really hard to spot, even by tools as *Zerogpt*.