Open Source Framework for Conducting Email Phishing Experiments

Gabriel de Melo Cruz
Faculté des Sciences et Techniques
Université de Bretagne Occidentale
6 Av. Victor Le Gorgeu
Gabriel.Demelocruz@etudiant.univ-brest.fr

***Abstract*— Social Engineering attacks are and have been a hot topic on cybersecurity for a long time now. Still, conducting robust and reliable experiments has proven to be challenging. In this work, we analyze current and past experiments in order to define an organized way to conduct email phishing experiments in the form of a study framework. We then implement a proof of concept experiment to illustrate a canonical use case of the framework.**

Keywords - social engineering, phishing, semantic attacks

I. INTRODUCTION

Social Engineering (SE) can be defined as the practice of "maneuvering human beings to take action in some aspect of their lives" [5]. Thus, Social Engineering is not, as it is frequently referred to, merely deceiving people in order to get access to information.

Social Engineering is also not necessarily linked to illegal activities. In fact, it is frequently used as a tool in almost every human knowledge area: advertisement techniques, economy stock manipulation and even children that cry to get what the candy they want from their parents, to name a few.

In this work we analyse the role of Social Engineering Attacks (SEAs) on cybersecurity, more specifically those defined as Semantic SEAs. We present and compare the most common types of SEAs. We identify current limitations of studying Social Engineering in the context of cybersecurity and we compare different state of the art experiments, their limitations and successes. Finally, we propose a framework to conduct practical and safe phishing experiments and build a proof of the concept architecture that can be easily used.

A. SOCIAL ENGINEERING ATTACKS

There are many definitions of Social Engineering Attacks (SEAs), it all depends on which definition of Social Engineering is assumed, and which kind of attacks are considered to be SE attacks. One of the most well-known SE attack models is the *ontological model*, in which a SE attack is defined as an attack that "employs either direct communication or indirect communication, and has a social engineer, a target, a medium, a goal, one or more compliance principles and one or more techniques". The ontological model also describes the main devices a SE attack contains [10]:

- **Social engineer:** the attacker (a person or an organization)
- **Target:** a person or an organization that the attacker is trying to get advantage of
- **Medium:** anything the attacker uses to get in contact with the target (i.e. email, social networks, SMS, paper mail, etc.)
- **Goal:** the objective of the attacker, this may be a piece of information, money, trust, etc.
- **Compliance principles:** psychological tools the attacker might use to engage the target (i.e. friendship, authority, charm, etc.)

B. SEMANTIC ATTACKS

As we have seen, the definition of Social Engineering Attacks is quite wide. This could lead us to some difficulties isolating the problems we want to tackle when researching. According to Heartfield and Loukas (2015):

> "This does not differentiate attacks that bypass the security of computer systems from those that can be observed in a nontechnical arena, such as prize-winning letter scams or physically impersonating an authority figure." [6]

Thus, in this work we will use the definition of *Semantic Social Engineering Attacks* instead. The term *Semantic Attack* is defined as:

> "The manipulation of user-computer interfacing with the purpose to breach a computer system's information security through user deception." [6]

Although it may not cover all types of SE attacks, using this definition allows us to focus on a more specific and, more importantly, more technical side of Social Engineering, one that is cybersecurity-focused.

C. SEAs in Cybersecurity

The Cybersecurity curriculum recommendation from the Association for Computer Machinery (ACM) cites the most important types Social Engineering attacks: phishing, spear phishing, physical/impersonation, vishing (phone phishing), email compromise and baiting. [1]

Phishing is "a form of deception in which an attacker attempts to fraudulently acquire sensitive information from a victim by impersonating a trustworthy entity". [16] A phishing attack could be as simple as an email which deceives students looking like it comes from a professor containing a link to a webpage with malicious code.

These attacks, however, can be way more sophisticated. An attacker who wishes to be more thorough could block a

user's Facebook password and then send him/her an email with a malicious link claiming to be Facebook with its password reset link. [16] Even though these two attacks have different levels of technical skills needed to be executed, fundamentally, both follow the same model.

Spear Phishing attacks, also called *context aware phishing attacks*, take advantage of other attributes specific of a given target. This kind of attack is much more effective but it generally takes much longer (weeks, months or even years) for the attacker to collect enough useful data from the target in order to claim to be some individual or organization much more convincingly.

In this work we are going to focus on experiments involving email mass phishing attacks, since these are one of the most common and potentially most dangerous types of SEAs.

II. SOCIAL ENGINEERING EXPERIMENTS

In this section we analyse well-known works that conduct social engineering experiments and evaluate the collected results. We point out each of their limitations to explain why our work is relevant.

A. On the anatomy of social engineering attacks — A literature-based dissection of successful attacks

In this paper, the authors collected and classified a handful of Social Engineering attacks from four different books. They try to dissect each of the attacks and get a grasp of the main steps and techniques each attack employs, and then they develop some statistics on top of that. [3] The problem with this research is that they takes the examples from books. There is no way to verify the validity of the stories, and there is no exact metric we can take from it. Although the books are very trustworthy, we are bound to the facts the authors decided to tell us, and, more importantly, *how* they decided to tell us.

B. Who Falls for Phish

On this more realistic study, the authors conduct an online research in phishing techniques with more than 1000 participants. They utilise and compare different types of phishing emails. [15] Finally, they compare the different levels of susceptibility according to demographic features.

Other than measuring the weaknesses in systems (and, here, in people), it is also useful to be able to measure the *resistance* to attacks that is, which attacks fail, why and in which contexts. In the article "Measuring Resistance to Social Engineering" the authors define a simple success metric, select a study population and conduct the attack. Then, they proceed to calculate some statistics about the studied group and the results obtained [4]. Also, "Who Falls for Phish" doesn't expose a crucial part of the work: the contracts and logistics of conducting such an attack. Without knowing the method used, it becomes extremely hard to estimate the reliability of the experiment and, more importantly, to replicate it.

C. Experimental Social Engineering Investigation and Prevention

Finally, this thesis proposes conducting an experiment with three different types of SE attacks: Face-to-face, email and telephone. As the scope of this work is Semantic SEAs, we are only interested in the latter. In the study, the author first compiles other articles from the literature into a table containing the information then known. He then proceeds to implement his own attacks and then compares the obtained results to those of past researches. [2] Again, this study lacks a very important piece information required in scientific research: method exposure for experiment replicability.

III. GOALS

The case studies done in the previous section show that the biggest limitations regarding Social Engineering experiments, and in particular Semantic Social Engineering experiments, are data disclosure and method reproducibility.

In the following sections we describe a general, practical method to conduct email phishing experiments. Our goal is to point out known limitations and problems researchers might encounter when first going into the field, as well as common mechanisms to overcome these limitations.

We then build a Proof of Concept free and open source framework to illustrate the use of our framework in a standard use case. Finally, we discuss limitations and possible future work that could be done to enhance the techniques here described.

It is important to point out that this is an example use case, and so other different contexts may require slightly different uses of the framework. Thus, the reader is encouraged not to take this work as an immutable set of rules, but rather as a rather flexible set of tips on how to conduct general Social Engineering research.

IV. FRAMEWORK ARCHITECTURE

The email phishing framework here developed can be broken down in four main steps: Consent, Bait, Result Analysis and Disclosure. Three of these steps, namely Consent, Bait and Disclosure, are interactions that directly involve the candidates of the experiment.

A. First Interaction: Consent

When working with people and deception, and especially including Social Engineering attacks, it is important to *always* have consent of some form. And, although consent *can* be given in verbal form, it is generally preferred, when dealing with scientific research and disclosure, to have written consent before conducting any kind of experiment. There are two general forms of consent: direct and indirect.

Direct consent is the most straightforward form of consent, where the individual him or herself gives their consent, either in verbal or written form. This may seem like the general overall form of consent, but the difference becomes clearer when indirect consent is explained.

Some types of contracts, in general at enterprises, universities or other organizations, often include specific clauses

that allow the organization to apply certain kinds of studies or exams without further direct consent from the individuals. These contracts illustrate have the **indirect consent** of the parties involved [12].

In the scope of this framework it is suggested to make use of indirect consent when possible, as it allows to conduct the experiment without immediate knowledge from the candidates, and thus yielding more reliable results.

On the other hand, indirect consent may not always be an option, and so researchers may recur to direct consent. In this case, an *informed consent term* must be signed by all the candidates. Depending on the context, this term may or may not allow the researchers to hide the subject of research at some degree from the candidates, which is advised in order to yield the most reliable results possible.

B. Second Interaction: Bait

This is the step where the attack is vehiculated. This framework is mainly focused on email phishing attacks, but adaptations can be made to enable other types of attack vectors as well.

There are two main characteristics of the attack that must be considered at this point: the *attack vector* and the *phishing method*. Each vector and method have their own tradeoffs. This does not mean that one vector is better than other, or that one method is worse than other, only that they are different. For the sake of conducting research, all of them are of equal value in the sense that they all will yield interesting results, regardless of the actual outcome of the attacks.

Perhaps the most straightforward, and thus the most common, attack vector is to send a unified, generic phishing email to all the targets at the same time. And, although this approach may be convenient from the attacker's point of view, it might not be the most effective in terms of getting targets to actually trust the bait. One alternative might be to send different sets of emails to different targets, or even send different phishing emails to the same target in order to compare effectiveness.

The second important aspect of baiting is the phishing mechanism used to gather information. This could as straightforward as a link to a clone website containing a fake form where users would submit their information to the attacker, or an email embedded with malicious javascript code that would expose the user's IP address or other sensitive information.

C. Result Analysis

This is the most flexible step of the framework. Here it is basically up to the researcher to decide which statistics or other comparison methods to use in order to analyse the data, according to the original thesis being tested. Then, it is also important to produce possible explanations for the results when they do not match the desired results.

D. Third Interaction: Disclosure

Data collection and disclosure are perhaps the most problematic steps of every social engineering research experiment. There is a multitude of laws and regulations that protect private data as well as copyright issues and other limitations. Also, every country has its own set of laws to cover data privacy, and so a very careful work has to be done when anonymizing data and designing contracts or other legal documents. [12] [7]

Nevertheless, disclosing the collected data as well as the methods used to collect it is a fundamental step when developing reliable, reproducible scientific research. It allows other researches to reproduce the experiment and compare results or even add more results to an existing dataset which could then be used to develop more robust mitigation techniques. This framework strongly encourages researchers to anonymize and disclose the collected data.

Social Engineering experiments may or may not be classified into educational or informational research. In this type of research, the disclosure, or *debriefing*, part of the research contains a step where researchers provide feedback to the candidates after the results are analyzed. Here, researchers can educate the candidates about best practices to stay safe online and not to fall in these types of attacks. This can be a very powerful incentive to convince organizations or candidates to apply to a research, which makes the whole process of getting consent much easier.

V. ETHICAL QUESTIONS

One of the most important questions that come to mind when researching malicious Social Engineering Attacks is: "Is manipulation always wrong?". The trivial, but rather naive, answer is "Yes, of course, one should never try to manipulate another human being". Perhaps this comes from the fact that we tend to think of manipulation in a more abstract way.

There are two main sub questions that compose the Ethical dilemma of manipulating others. The first is *evaluation question*: "How should we evaluate the moral status of manipulation?". The second can be worded as "How can we identify which forms of influence are manipulative and which are not?".[11] This last question is called the *identification question*.

Researchers need to be aware of the consequences their experiments, especially the most sensitive ones like SEA experiments, might have on the target audience. Although it may not be the purpose of this work to solve or dive deeply into philosophical questions, pointing them out before everything else is a crucial part of any social engineering experiments.

VI. PROOF OF CONCEPT

In order to give the reader a practical example of how to use the phishing framework here described, we build a proof of concept attack that implements the most important parts of the framework in their respective canonical forms, that is, the most common use case of the framework.

The proof of concept intends to conduct an experiment with the students of the School of Sciences and Techniques at the University of Western Brittany (UBO in french). We create a clone website with a malicious form that sends the

information typed in to the attackers. We also craft emails to be sent in a way to the clients.

It is important to point out that if need be the use case can and should be slightly different than the canonical one, but overall the proof of concept should be enough to give the reader a general guide on how to implement the attacks.

A. Targets / Subjects

This experiment intends to target undergraduate and graduate students of the School of Sciences and Techniques at UBO.

B. Technologies

The main technologies used to build the proof of concept were:

- **Google Sheets:** Cloud-hosted spreadsheets platform, used to store collected data from the candidates. [9]
- **Google Forms:** Cloud-hosted form platform, used to issue the first consent form and get the consent from the candidates. [8]
- **wget:** Command line tool for retrieving files using HTTP, HTTPS, FTP and FTPS. Used to clone the website. [13]
- **Flask:** Microframework for building web applications. Used to host and deploy the clone website. [14]
- **Email Server:** UBO's SMTP email server. Used to send the bait and the disclosure emails.

C. Consent

The first step described by the framework, and perhaps the most important and sensitive one, is *consent*. We issued an email to all the undergraduate and graduate students of the School of Sciences and Techniques inviting people interested to participate in the experiment (see Figure 1).

> Hello,
>
> The Internet is growing fast, and with it come security risks. In this context, the student Gabriel de Melo Cruz, working with the professors Karim Bigou, Catherine Dezan and David Espes, is conducting a study to identify the key factors that make users vulnerable. To participate in our study, register by filling out *this form*.
>
> Regards

Fig. 1. Email inviting candidates to participate on the experiment.

This email contains a link to the subscription form, which included a direct consent contract that specified the scope of the experiment without fully disclose the means or the specific ways used to deceive the candidates (see Figure 2).

Although this is not the ideal scenario, because candidates would already be biased, it was the chosen consent method because the university didn't have a contract that allowed otherwise.

> **Online scam experiment**
>
> Thank you for your interest. This experience will test your ability to identify online scamming techniques. After filling out the form, we will try to apply certain deception techniques on you over the next few weeks. Due to the nature of this experience, we cannot disclose exactly what techniques we intend to use, but we guarantee that no one will attempt to contact you physically (i.e. in person).
> By clicking on "I accept the conditions described above", you agree that we perform online scamming deception techniques in you and that we publicly disclose the data collected herein, excluding your email address mail and your name - these will only be used to get in touch with you and nothing more during or after the end of the search.

Fig. 2. Form containing the consent contract information.

D. Bait

With the target emails collected and consent form signed, we would then proceed to prepare the attack: a simple clone of the university's website containing a malicious form (see Figure 3), and finally send the bait emails (see Figure 4).

Fig. 3. Cloned website with malicious form.

E. Disclosure

After compiling and analysing the results, data and experiment disclosure is extremely important. In this theoretical experiment, we send feedback emails to the candidates explaining the malicious tips they failed to identify. Figure 5 shows an example feedback email.

This step if crucial for two main reasons. The first is educating the candidates, since this particular example has also an educational side to it. It is important not to compare

Hello,

We have created an email list to send official university information on the pandemic, we ask everyone to subscribe on *this form*.

Regards

Fig. 4. Email to bait targets into clicking a malicious link.

candidates, as this could cause distress and other undesired feelings especially in more vulnerable candidates.

The second main reason why disclosure is important is because it is in this step that the researcher has the chance to expose their methods and the collected, and possibly anonymized, data.

Hello,

Here are the results of your participation in the online scamming experiment:

- You clicked on a malicious link: TRUE
- You have completed a form with your contact information: TRUE
- You have completed a form with your course information: FALSE
- You have completed a form with your personal information: FALSE

Regards

Fig. 5. Feedback email containing the results of the experiment of a particular candidate.

VII. LIMITATIONS AND FUTURE WORK

A. Scope

The biggest limitation this work currently faces is scope. This framework is designed to a very particular use case: email phishing experiments. Even though email phishing attacks are very common, which may compensate for the narrow scope, more work is needed testing the framework in other scopes, such as baiting phone calls and other types of semantic attacks, in order to identify its limitations and strengths.

B. Tests

Testing is a very intrinsicate part of theoretical work. As this is an organizational framework on conducting phishing techniques, it is important to test it on real experiments and scope out its limitations, add more details to it and make it more robust in general.

C. Subjectivity

Finally, as it is with all scientific work, this framework is subject to individual subjectivity. It may be a good guide to some researchers that are beginning on the field and need some general guidance or even experienced ones that need a formal structure to conduct several experiments. However it may also not be ideal for all use cases.

VIII. CONCLUSIONS

In this work we discussed the role and definition of Social Engineering and Social Engineering Attacks, as well as the subset of Semantic Social Engineering Attacks. We then analyzed current and past Semantic Social Engineering experiments and pointed out their limitations.

With that in mind, we then defined an organized way to conduct an email phishing experiment in the form of a study framework. Finally, we implemented an example experiment to illustrate a possible frequent use case of the framework.

Finally, we discussed the current limitations with the actual implementation and the necessity of more tests and actual field experiments to figure out the limitations and strengths of the framework, as well as to point out the use cases where it does not apply very well.

ACKNOWLEDGEMENTS

Special thanks to Catherine Dezan, David Espes and Karim Bigou, for the countless hours of guidance, reviews, comments and tips during the entire research process and also for giving me the opportunity of meet France for the first time. The learning experiences of being in touch with the culture and other aspects of french life definitely reflected in this project.

REFERENCES

[1] ACM. “Curriculum Guidelines for Post-Secondary Degree Programs in Cybersecurity”. In: *A Report in the Computing Curricula Series Joint Task Force on Cybersecurity Education*. Ed. by Edward N. Zalta. 31 December 2017. Associataion for Computer Machinery, 2017.

[2] Jan-Willem Bullee. “Experimental social engineering: investigation and prevention”. English. CTIT PhD Thesis Series No 17-443, ISSN 1381-3617. PhD thesis. Netherlands, Oct. 2017. ISBN: 978-90-365-4397-2. DOI: 10.3990/1.9789036543972.

[3] Jan-Willem Hendrik Bullee et al. “On the anatomy of social engineering attacks—A literature-based dissection of successful attacks”. In: *Journal of Investigative Psychology and Offender Profiling* 15.1 (2018), pp. 20–45. DOI: 10.1002/jip.1482. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jip.1482. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jip.1482.

[4] Jan-Willem Bullee et al. “The persuasion and security awareness experiment: Reducing the success of social engineering attacks.” In: *Journal of Experimental Criminology* 11 (Mar. 2015). DOI: 10.1007/s11292-014-9222-7.

[5] Christopher Hadnagy. *Social engineering: the art of human hacking*. Wiley, 2018.

[6] R. Heartfield and G. Loukas. "A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks". English. In: *ACM Computing Surveys* 48.3 (2015). Cited By :40. URL: www.scopus.com.

[7] Kaspar Jüristo. "How to Conduct Email Phishing Experiments". English. PhD thesis. 2018.

[8] Google LLC. *Google Forms*. URL: https://www.google.com/forms/about/ (visited on 07/01/2020).

[9] Google LLC. *Google Sheets*. URL: https://www.google.com/sheets/about/ (visited on 07/01/2020).

[10] F. Mouton, L. Leenen, and H. S. Venter. "Social engineering attack examples, templates and scenarios". English. In: *Computers and Security* 59 (2016), pp. 186–209.

[11] Robert Noggle. "The Ethics of Manipulation". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2018. Metaphysics Research Lab, Stanford University, 2018.

[12] Markus Jakobsson Peter Finn. "Designing and Conducting Phishing Experiments". In: (). URL: http://www.icir.org/vern/cs261n/papers/finn-jakobsson-phishing-design.pdf.

[13] GNU Project. *Wget*. URL: https://www.gnu.org/software/wget/ (visited on 07/01/2020).

[14] Pallets Projects. *Flask Web Microframework*. URL: https://palletsprojects.com/p/flask/ (visited on 07/01/2020).

[15] Steve Sheng et al. "Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, pp. 373–382. ISBN: 9781605589299. DOI: 10.1145/1753326.1753383. URL: https://doi.org/10.1145/1753326.1753383.

[16] Markus Jakobson Tom N. Jagatic Nathaniel A. Johnson and Filippo Menczer. "Social Phishing". In: *Communications of the ACM* 50.10 (2007).