Automated Social Engineering Attacks using ChatBots on Professional Social Networks

Maurício Ariza[1], **Antonio João Gonçalves de Azambuja**[1], **Jéferson Campos Nobre**[1], **Lisandro Zambenedetti Granville**[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{mariza,antonio.azambuja,jcnobre,granville}@inf.ufrgs.br

Abstract. *The growth of the internet and social networks has intensified human interactions, raising the risk of cyberattacks. Social Engineering targets those human relationships in the cyber environment, using technology as a support to exploit natural human failures. Research has shown the capacity of Social Engineering attacks, however, there are few papers focusing on the evolution and trust of ChatBots and automation as a support for those attacks. This paper presents an analysis of the capacity of professional social networks to detect and block automated Social Engineering threats to their users. The approach developed allowed us to identify the characteristics of the trust relationship between the user, the social network, and the ChatBot resulting from the established interaction, and failures on the part of social networks to identify and block this kind of behavior. To this end, an automated Social Engineering bot was developed. The analysis and discussion of the results allow demonstration of the security vulnerabilities present in professional networks and in building the user's trust relationship with the ChatBot.*

1. Introduction

Social networks have been used as a vector for cyber attacks to obtain sensitive information from users using virtual profiles [Paradise et al. 2019]. Attackers make use of the connectivity of these social networks to expand their area of operation, a fact that exponentially increases the challenges of cybersecurity. The interconnectivity of social networks and the growth of the cognitive dimension of work are making human resources one of the pillars of security [Culot et al. 2019] [Greitzer et al. 2019].

Cyber attacks carried out on social networks have exploited human interaction in conjunction with technological gaps, weakening the cybersecurity chain. Organizations have used defense solutions to face cyber attacks, such as firewalls, Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), and antivirus. However, these defense mechanisms have not been sufficient to fully prevent Social Engineering (SE) actions in the cyber environment [Salahdine and Kaabouch 2019] [Klimburg-Witjes and Wentland 2021].

Cybersecurity experts characterize attacks that focus on human behavior as SE attacks that aim to manipulate users to reveal sensitive information. These attacks combine human interaction with the exploitation of [Klimburg-Witjes and Wentland 2021] technological vulnerabilities. The increasing use of social networks to establish personal and

professional relationships opens a field for the actions of Automated Social Engineering (ASE) bots. [Huber et al. 2009]. Social Engineers have sought to develop bots with intelligence, making automated interaction unnoticed by users.

Bots can be used for positive actions, such as helping the user in their online experience. However, bots developed for ASE attacks have made it possible for a single attacker to contact a large number of potential victims simultaneously because of their scalability. The attacker aims to get the victim to reveal sensitive information, which can be used for data theft [Huber et al. 2009] [Dewangan and Kaushal 2016]. Bots are automated software, which sometimes uses features such as artificial intelligence. In its functionality it has the ability to execute operation and control commands to impersonate humans, simulating the activities of real users [Shafahi et al. 2016].

During the literature review, few papers were identified that present analyses on ASE with the use of bots. The studies focus on human behavior in the face of SE actions [Huber et al. 2009]. In this sense, the strategies of a cyber attack are being framed as a social issue and not just a technical vulnerability, according to the authors [Klimburg-Witjes and Wentland 2021].

SE attacks using fake accounts with identity theft on social networks with impacts on users' privacy and information security were already analyzed [Al-Charchafchi et al. 2019]. The discussion of the use of SocialBots for social media conviction campaigns is present in the papers that evaluate the impact of these tools on user behavior [Boshmaf et al. 2013]. The use of bots to influence users of Twitter, aiming to gain followers and compromise the network structure are presented in the work of some authors [Freitas et al. 2015] and [Messias et al. 2018].

An analysis of SE tasks automation through a bot on Facebook [Huber et al. 2009] concludes that persuasion is an essential resource in the ASE process. However, no papers were identified that present bots with intelligence to perform an automated human interaction to search for sensitive information, without the perception of the user being targeted by the ASE technique.

This paper proposes a proof-of-concept bot with intelligence to perform an ASE attack, having the offer of attractive jobs as a stimulus for user interaction with the bot. This attack focuses on a specific group of users of the professional social network LinkedIn. The main contributions of this work are: (i) evaluate LinkedIn capacity to detect and mitigate an ASE attack; (ii) implement a proof of concept to validate the technical viability of this attack; and (iii) propose control improvements that can be implemented by those social networks to decrease the risk of SE attacks to their users.

The paper is organized as follows. In Section 2 the concepts related to the theory to support the research. Next, Section 3, presents the methodological proposal of the study and the identified limitations. Section 4, performs an evaluation of the proposal and a discussion of the results. Section 5, analyzes the related works. Finally, it presents in Section 6 the conclusion and an approach for future work.

2. Background

2.1. Bots

In cyberspace, there are authentic bots that aim to perform useful activities for users. However, there are also malicious bots, which can perform attacks to obtain relevant information or maintain control of the accessed device. Bots can be used for spreading false information (fake news), spam, and phishing. [Freitas et al. 2015].

Cybercriminals use malicious bots to simulate human behavior, bypassing security mechanisms. As mentioned by some authors [Huber et al. 2009], ASE attacks using bots take SE to a new level of scalability of attacks. In the context of ASE, cybercriminals use malicious bots to simulate human behavior, avoiding security mechanisms.

With the growth of social networks and the large volume of data in cyberspace, social engineers have started to spread bots with human-like behavior to a large number of users. These bots simulate human conversations, known as ChatBots and, those that operate on social networks, SocialBots [Shafahi et al. 2016].

SocialBots are defined as effective tools to perform SE attacks, with the aim of gaining access to sensitive information. It is a tool that has the ability to compromise the structure of social networks, aiming to: (i) steal identity; (ii) influence users; (iii) increase the number of followers; and (iv) inflate the popularity ratings of a particular profile account [Boshmaf et al. 2011] [Camisani-Calzolari 2012] [Dewangan and Kaushal 2016].

As such, SocialBots need a technical infrastructure, with a combination of a social networking platform and technical requirements for automating the behavior of an account, using an Application Programming Interface (API) or proprietary mechanisms to interact with the platform [Assenmacher et al. 2020]. As a certain degree of intelligence is incorporated into SocialBots to simulate human behavior, it sparks interest in research on the topic [Ferrara et al. 2016].

It is a tool that simulates human behavior to perform automated interactions on social networks [Rouse 2013]. SocialBots for the most part are automated social media accounts that impersonate people. These interactions are artificial intelligence activities that have shown growth in the online environment with the use of this tool [Freitas et al. 2015] [Hepp 2020].

ChatBots are the integration of systems, tools, and scripts that promote instant messaging conversations with or without human participation [Stoeckli et al. 2018]. They are developed to help human users in specific service situations and are not exhaustive. Here are three (3) examples: customer service, communication service, and digital education service [Grimme et al. 2017].

This tool originated in the field of Computer Science, in this sense it is a tool developed with the help of artificial intelligence mechanisms that interact with users. The use of natural language in ChatBots, a language used for human communication, is a challenge to be overcome for the development of the tool [Khan and Das 2018].

The growing use of personal assistants demonstrates the popularity of ChatBots. However, as the use of this tool grows, it is important to keep in mind the increase of attacks on typical ChatBots architectures, for example: client module, communication module, response generation module, and database [Ye and Li 2020].

These tools, SocialBots and ChatBots, have been developed with the help of artificial intelligence mechanisms that interact with users [Freitas et al. 2015]. Artificial intelligence is similar to human intelligence, developed with automation as per the need of the application [Ferrara et al. 2016]. As a certain degree of intelligence is built into the tools to simulate human behavior, the capacity and scalability of attacks increase.

2.2. Automated Social Engineering

Social engineers make use of automated Bots, which are able to impersonate humans to carry out an ASE [Shafahi et al. 2016]. These attacks seek to establish a trust relationship to obtain sensitive information about the user and require little intervention to establish the relationship, enabling greater reach by their scalability [Mitnick and Simon 2003] [Huber et al. 2009]. Human communication has been based on the development of human-machine interfaces. The disruptive technologies are inspiring studies on this communication [Guzman and Lewis 2020].

Social networks are facilitating communication, social interaction, and sharing of personal and corporate information, increasing their popularity in the cyber environment. These networks represent an attractive virtual space for attackers to exploit technical vulnerabilities and users' lack of knowledge and awareness of SE actions [Al-Charchafchi et al. 2019].

The relationships formed in this environment allow greater exchange of information, ratifying the statement of [Castells 2009], that in social relationships, networks are communicative structures. Cyberspace constitutes a promising scenario for the practice of all sorts of illicit acts, without respecting geopolitical borders. The growth of social networks has enabled the creation of a large number of fake profiles, with the use of automated Bots for this activity [Tiwari 2017].

It is common the SE attacks to require time to establish a trusting relationship and resources. However, SE can be accomplished through automated mechanisms. ASE attacks require little human intervention to establish the relationship and have greater reach because of their scalability [Huber et al. 2009]. Automated attacks can be prepared using valuable information and/or influencing certain groups in social networks [Gallegos-Segovia et al. 2017]. ASE attacks using Bots and Phishing have become more frequent due to the increasing use of social networks for personal and professional activities [De Kimpe et al. 2020].

3. Methodology

It is already a common concept to classify humans/users as the weakest link of the cybersecurity chain, as attacks exploiting their failures have better success rates and sometimes require less technical skills and risk for the attacker [Darwish et al. 2012].

Social networks have their essence based on creating interaction between humans. But beyond allowing a space where distance boundaries can be bypassed to enhance connections, they also bring to the virtual world many of the threads from the real world. But there is a difference as it is a space where people don't have the same awareness and capacity to recognize risks, which together with the stronger capacity of anonymity and impersonation become a perfect environment for SE [Crossler and Bélanger 2014].

Comparing to other social networks like Facebook, Twitter, and Instagram, professional social networks create a more corporate environment, focused on business connections and career growth. This scenario creates a sense of trust and credibility, attracting headhunters looking for candidates as well as companies looking for potential new customers. These relations are already exploited by social engineers, especially impersonating recruiters using attractive job opportunities as bait to steal internal information or personal data of the victims[1].

Currently, LinkedIn is the most popular professional social network, with more than 850 million members in more than 200 countries. Their User Agreement defines in section 8.2[2] the actions that are not allowed to users, highlighting forbidden use of false information or impersonation in the profile and usage of bots and automation to realize actions in the platform.

Considering the SE attacks already mentioned, we can easily find references to fake profiles or false information used for several reasons. Talking specifically about automation, if we search on the internet or code repositories like GitHub we could find several bots and scripts specifically designed for LinkedIn. Those references indicate a potential lack or insufficient implementation of controls by the platform, that looks to focus the enforcement of policies on complaints or reports done by the users. Our goal is to evaluate the risk level for LinkedIn users to face an ASE attack, understand the platform's capacity to detect and block those attacks, and offer suggestions to improve their controls to decrease risk with no or minimal impact on usability.

3.1. Limitations

SE was born in the field of psychology, as they aim to exploit human failures, using technology as a support to achieve those goals. For a full understanding of the impact of a SE attack, we will need to only validate technical aspects, but also tricky subjects to observe their behavior and actions. The field of social psychology explored those matters for a while to could identify which are the boundaries for ethical research, understanding that the goals do not justify the methods to avoid extreme cases like the famous Milgram experiments in the 70s.

Exposing people to situations where they will be deceived, and having their vulnerabilities exploited without their consent (or full understanding) violates ethical dilemmas. As result, later they can create frustration and stress due to expectations created, broken promises, or the feeling of being fooled. Understanding and respecting those boundaries was one of the drivers of this paper, and even as discussing research ethics was not our goal it's not possible to run a work like that without raising questions in this matter.

Our first challenge was how to validate our proposal within keeping compliance with ethical policies. In order to achieve that, we break the entire attack life cycle in separated steps and try to do testing and validation of each one individually, based on the results we could have a fair understanding of the application response and the potential of the full attack.

[1] https://www.forbes.com/sites/reneemorad/2017/06/30/how-to-avoid-the-latest-linkedin-scam/?sh=13e1d13849c1

[2] https://www.linkedin.com/legal/user-agreement#dos

The main difficulties happened in the Approach and Interview steps, as they would require at least some level of contact with subjects. For the Approach step, we achieved a tiny line between keeping our premises and violating ethical barriers. We then decided to limit to a minimum number the quantity of LinkedIn users receiving the request and the message. In this way, we could evaluate if the platform will identify the automated behavior, and then cancel/exclude all actions immediately as a damage control mechanism. This format allows us to avoid any individual really having contact with our testing accounts.

For the Interview step, we focused on the main functionality of our Recruiter Chatbot: do a job interview. As the malicious action would happen by making the victim believe it is a real job interview happening, and them being expected to have a process involving signing a contract and sending documents for identification, for example. So our tests tried to validate the bot's capacity to run a convincing job interview as a way to consider their potential to have the same results in a full cycle attack.

Considering all those mechanisms we evaluate as being achieved enough results to validate the potential of the proposed attack without violating any associated ethical requirements. We believe this is a core topic for any kind of research and a deeper analysis of the impact of ethical research matters especially on the SE field it's an intriguing topic for further studies.

4. Proposal and Evaluation

To evaluate the attack we used a proof of concept scenario with 2 bots. The first one interacts with the social network to search and contact the victims with the bait - the Platform Bot. The second one it's a Chat Bot service that would act directly with the victims to execute the step of the job interview - the Recruiter Bot.

For the Platform Bot role, we developed a Python code to connect with LinkedIn. LinkedIn offers a very rich API for software interaction, but considering the characteristics of the attack and the kind of validation we are looking for it would not be the best option. Then we evaluate that a connection is done through a browser - like done by any regular user - would provide us a better understanding of the social network response than a channel for software connections. In order to achieve this we use the Selenium library, which allowed our bot to act in a request-response through the browser.

For the Recruiter Bot, there were several options available that could fit into the need to execute the necessary actions without the need to develop custom code. Using a pre-defined set of job interview questions, plus information scrapped from the victim's LinkedIn profile, the Recruiter Bot would basically conduct a false job interview with the victim with the goal to collect sensitive information from current and past jobs. As it can execute both RH-like interviews with more generic questions and a technical interview, the entire process can be executed by the same bot and then, in the end, as the victim is 'accepted' for the position, personal information can then be stolen for identity theft to sign the fake work contract.

Our attack proposal follows the SE attack stages structure [Mitnick and Simon 2003], organizing it in 4 steps: i) Authentication, ii) Search, iii) Approach, and iv) Interview. Our goals in each are detailed below:

1. Authentication: As also illustrated in Figure 1, the goal of this step is to verify if the social network detects or has different behaviors when the user logon process it's done using automation. For this evaluation, our Platform Bot opens the LinkedIn website in the browser, maps the source code of the main page to identify the credential fields, fills them with the values received, and then submits to conclude the authentication process and access the main page of a logged user.

2. Search: This step aims to check the detection of automated searching of users. Similarly to the first step, the Platform Bot maps the page source code, identifies the search field, runs the search for the provided terms and then stores temporarily the returned profiles, creating a database of potential victims. Figure 2 also refers to this step.

3. Approach: Goal here is to start the interaction with the profiles of potential victims collected in the previous step. As also seen in Figure 3, using the created temporary database, the Platform Bot adds them as contacts and sends a custom message, which serves as the bait for interaction. This action happens to all profiles captured.

4. Interview: Based on the results of the bait sent in step 3, a script scrapes data from the victim's LinkedIn profile to feed the Recruiter Bot database, which they have enough information to execute a job interview with the victim. Figure 4 also illustrates the full cycle of this step.

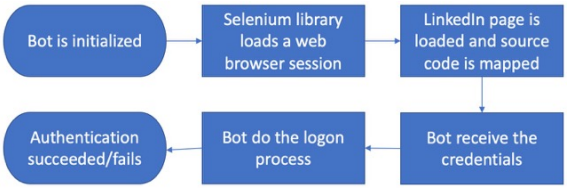

Figure 1. Attack authentication phase.

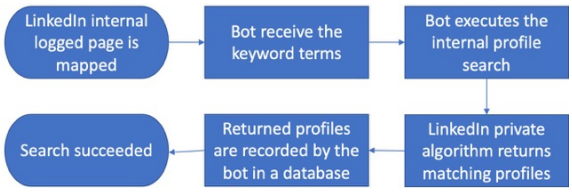

Figure 2. Attack search phase.

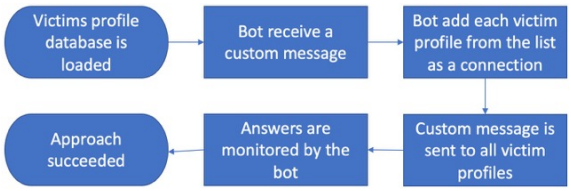

Figure 3. Attack approach phase.

4.1. Testing

The testing phase followed the steps of our proposed Attack Flow. The Platform bot was executed in a Windows machine running Python, the Selenium library, and Google Chrome as a browser. For the Interview bot we used the SAP Conversational AI platform.

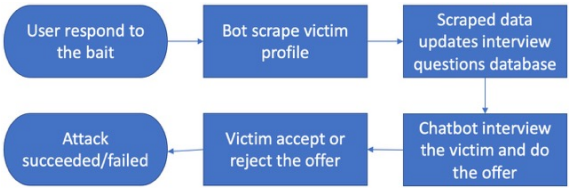

Figure 4. Attack interview phase.

Testing Step 1 - Authentication: The success criteria of this step is to execute authentication in the platform following different behaviors to observe if we have any impact from the application side or that could demonstrate controls or blockers due to automation characteristics. For comparison criteria, we defined three basic behavior patterns to be tested: (1) Do the logon process 10 times simultaneously, (2) Do the logon process 10 times with 5 seconds waiting time between each attempt, and (3) Do the logon process 10 times with 10 seconds waiting time between each attempt. Those patterns try to replicate behaviors not expected from a real human user due to the quantity or speed of attempts, especially as the execution is happening through the web browser. Also, for each one of them, we tested using the following variations to observe if they impact the results in any way:

- Receive the credentials of the created fake profile in execution time through the script.
- Read the credentials of the created fake profile from a file.
- Use of wrong/invalid credentials.
- Use of a public proxy to execute the logon from a random country, different from the one defined as the location of the user in the created fake profile.

In the results of our tests, we did not observe any difference in the social network behavior when the tests were executed using the valid credentials of the fake account. We also executed the test of each pattern on different days to guarantee that the execution of one of them would impact the results of the others. As an alternative variation, we also executed 10 sequential login attempts using the valid fake profile credentials and invalid ones, but manually (not using the bot), through a web browser, which did not demonstrate as well any difference in the results. When testing using invalid credentials, both through a bot or manually, after the 6th attempt LinkedIn starts requiring a puzzle (similar to a Captcha verification) and/or additional validation like a code sent by email/SMS to proceed with login, indicating that brute force behaviors are identified and blocked, which not happens for all kind of automated access attempts.

Testing Step 2 - Search: In this step we evaluated the capacity of the Platform Bot to execute queries in the social network without being detected. The authors defined a series of keywords to be used for the queries, based on some standard IT skills associated with this project only for reference. It is important to highlight that the quantity, order, and all other information related to query results are all associated with the search algorithm used by LinkedIn, which analysis is out of the scope of this work. Keywords used were only a manner to evaluate the response for automated queries through the web browser. For the test, we enumerate ten keywords, ”test”, ”Social Engineering”, ”Bot”, ”Chatbots”, ”Social Networks”, ”Information Security”, ”Python”, ”Automation”, ”GitHub” and ”API”. Similar to what was done in step 1, we used the following variations:

- Querying the same keyword 10 times simultaneously.
- Querying the same keyword 10 times with 5 seconds waiting time between each.
- Querying 10 simultaneous sessions, each one using a different keyword.
- Querying the 10 different keywords in the same session, in sequence, with 5 seconds waiting time between each.

Was not a goal of this step to do stress/load testing or cause a denial of service in the application. The tested behaviors used a speed and/or quantity not expected to be executed by a human user, especially as they were executed through the web browser. Besides the expected differences in the results (considering the different terms used and the LinkedIn algorithm, out of the scope of this paper), we did not observe any differences in the variations, and all queries received correctly the results with a list of profiles associated with the keyword term.

Testing Step 3 - Approach: In the previous step, after running the queries, the Platform Bot keeps a reference of the returned profiles to be used for this step. Here we had our main challenge on the already discussed ethical implications. Looking to have a limit on the impact of our research without prejudice to the results, we implemented the following controls to our bot:

- For each query, instead of the several pages of results, we keep only the ones on the first page, which were around 15-21 profiles per query.
- We only executed the approach 10 times, one per keyword, disregarding variations on queries.
- After executing the approach, the bot keeps a record of the success and then deletes/cancels all their actions within the user/victim.

The approach happened with a connection request to the user and the sending of a custom message, using tags to use the real user name instead of generic terms like 'dear user'. Again, once validated the request and the message, the request was canceled and the message was deleted for both sides, avoiding any further interaction with the users. Again no impact or actions from the side of a social network were identified during any of the tests.

Testing Step 4 - Interview: For this step our validation followed a different direction. We used the SAP Conversational IA platform, with a proof of concept chatbot to work as the Recruiter Bot. Based on an initial database of common questions for a job interview, we used a script to scrape the data from the victim profile and use them as input for additional questions, creating questions like "How was your experience in Company X?", "Can you talk more about your skills on technology Y?". Based on the observation of the chatbot interview for different profiles randomly selected from the previous step, it was capable to conduct a job interview without the need for management or additional command/control. This result allows us to demonstrate their capacity to be used for the proposed attack without having a realistic approach with subjects/victims, violating the already discussed ethical limitations.

4.2. Evaluation of Results

The main contribution of this paper was to look to validate the hypothesis of lack or insufficient controls implemented by social networks, no matter whether this could be known or expected in the technology field, we could not found official research or academic

references as foundations. It was not the goal to explore the human factor and the psychological matters associated with SE, but to observe if the technology channels allow or at least do not offer barriers to avoid those aspects being exploited in their users, especially in cases like our proposal were it's possible to achieve high scalability by the attacker.

Certainly the main impact of rigid controls on a social network it's on the user experience, which can directly affect the usage, base of users, and several important success indicators in this market. As a result, users can migrate to concurrent platforms, for example. But should be possible to find a balance and increase the security levels with no or minimal impact on the users.

Considering that the current LinkedIn User Agreement already forbidden the usage of automation, some automation-detection controls can be applied. This will not only avoid ASE but an entire behavior that is already forbidden. But even this control can be implemented with some flexibility. A regular user connecting through a web browser has some human limitations on their speed and quantity of requests - below a certain limit, not only you have a certain or potential automated behavior, but you also have high chances of SPAM and other unsolicited interactions. Based on our results, some examples of simple controls to detect automated behavior are:

- More than one simultaneous login of the same user (with some variations, like if you consider a user logged on the laptop and the smartphone at the same time), or several successful logins in a short time period.
- Several simultaneous and/or continuous requests (not only search queries but for any action) in a quantity or time frame higher than the average capacity of an human being.
- Do several contact requests and/or send several messages to different users simultaneously or in a short time frame (also potentially indicating SPAM).

The enforcement of controls on those behaviors doesn't need to be the block or cancellation of the request. Requiring additional fields like a Captcha, similar to what is already used to avoid brute force attacks, can be an excellent way to avoid automation, as they would be required only for certain scenarios that will not affect most of the regular users.

Going a bit beyond, imagining the need of certain users/scenarios where some automation can be useful or required - for real recruiters, for example, the enforcement of controls can be more rigid through a web browser (where regular users, not automation, is expected) and more flexible through API, for example. This will allow better monitoring and control by the platform, even being able from a business perspective to offer a certain quantity of free requests for minor customers (like independent small headhunters) or more robust professional services sold by the platform, like the already existing LinkedIn Recruiter.

Those examples of controls can solve the automation issue, enforcing already existing policies with minimum impact on users. But a more complex challenge is how easy is to create fake profiles, a problem not only for LinkedIn but for any social network in the current days. It is not difficult to build a base of interactions and contacts that can create a sense of legitimacy, organically through real users or even through a network of other false profiles.

There is no easy answer to doing it without complex validations. But the impact

of fake profiles has been growing so fast that discussions over mandatory user validation are already happening on other social networks like Twitter. Of course, those networks have a different set of users and characteristics of usage, but if we analyze that LinkedIn's mission is to "connect the world's professionals to make them more productive and successful"[3], would not be credibility and veracity of users a matter of interest for all their users? Verified profiles already exist usually for social influencers, and potentially even without the enforcement, many users will potentially look for this validation as a way to recognize their work and responsibility - or at least some groups like recruiters can be targeted. There are several opportunities, each with pros and cons, but certainly, some kind of control in this direction will be necessary to make it at least a bit harder to personify attacks.

5. Related Work

[Boshmaf et al. 2013], evaluate the vulnerabilities of social networks arising from a large-scale infiltration campaign using SocialBots. This study presents in their results an infiltration success rate of 80 % on Facebook, an index that demonstrates an unauthorized disclosure of private user data.

[Dewangan and Kaushal 2016], presents a model for detecting SocialBots used in political campaigns and marketing of products, having as input the behavior analysis. These actions bring with them security risks, considering the use of social networks for disseminating political positions and monitoring the consumption profile of users.

[Aroyo et al. 2018], discuss how SE exploits the trust relationship between users and bots. Based on the four (4) stages of an SE attack [Mitnick and Simon 2003], a bot was developed to simulate this task. First, the bot sought to obtain information with private questions. Then, it established a relationship of trust with the users, for a virtual and anonymous approach to the target.

With these actions, authors present in the research results that users have established a trust relationship with the tool. Among the requirements in the interaction with users, the ethical aspects were considered, by these authors.

[Al-Charchafchi et al. 2019], present a review of research on privacy and threats in social networks. For the authors, although the literature presents work on privacy, more effort is needed. The social networking environment is a rich source of personal data, making it an attraction for actions in social engineers, who exploit the users' lack of awareness and knowledge on security-related issues.

The complexity of SE attacks is related to the combination of social strategies and techniques used to carry out a cybercrime [Al-Charchafchi et al. 2019]. In this context to mitigate the impacts of attacks, [Piovesan et al. 2019] claims that security policies can provide a higher level of information security. However, they do not guarantee complete security.

[Freitas et al. 2014], present a discussion on the impact of the use of SocialBots on Twitter to characterize the behavior of the tool on a large database. In the results, the authors highlight that the method they developed to characterize and detect SocialBots, had a 92% successful detection indicator.

[3]https://about.linkedin.com

[Messias et al. 2018], claim that a simple Bot can achieve high levels of influence on Twitter. [Shafahi et al. 2016], on the other hand, points to the need to raise the level of awareness about phishing actions that use SocialBots. The authors state that these actions pose a threat to organizations.

[Paradise et al. 2019], analyze in the organizational context the strategies to monitor organizational social networks and detect SocialBots that aim to obtain data from the organization. The strategies were analyzed considering different levels of attacker knowledge using a simulation with real social network data.

[Huber et al. 2009], present the cycle of an ASE attack using a Bot. The attack demonstrated how social networks can be used by social engineers to obtain information. To this end, two (2) experiments were conducted in the study. The first analyzed the ability of Bots to obtain information from social networks. The second performed the Turing test, which seeks to evaluate the ability of a machine to imitate a human being.

Finally, for the authors, ASE with Bots is scalable and requires fewer human resources. The tool was used in a proof of concept on Facebook. The two (2) experiments allowed to ratify that it is possible to automate SE actions to obtain information and to demonstrate that the Bot used was not identified by the security measures of Facebook. The increasing number of users' social interactions on networks makes SE automation Bots an interesting tool for social engineers.

6. Conclusion

Cyber attacks have been exposing the vulnerabilities of computer networks and applications. Especially the context of social networks, becoming each day more important in people's lives, are being a promising scenario for several malicious actions, and the current defense mechanisms are not being efficient to mitigate or avoid them, highlighting the exploitation of trust using bots.

ASE bots offer great scalability with no need for more exposure from the attacker. This paper presented the development of a bot-based approach to simulate ASE attacks using job proposals as bait. Through a fake recruiter profile on LinkedIn, is it possible to identify and contact potential victims using automated mechanisms, looking for leakage of personal or corporate data. The complete absence of controls demonstrates the potential for similar SE actions in the real world, which should raise awareness and important actions to address those threats in an Information Security strategy.

The main contributions of this research are: i) Implement a Proof of Concept to validate the technical viability of this attack; ii) Evaluate the defense and response mechanisms of the social network to an ASE attack; and iii) Offer some potential ways to mitigate those attacks with minimum impact to user experience.

As a future work, the implementation of improvements to the proof of concept Platform Bot would allow to map and evaluate the limits for automated activities supported by the platform, and a more realistic simulation of an automated attack end-to-end. Also, more testing over the Recruiter Bot using real subjects outside an attack scenario will also help to better understand the capacity of similar chatbots to convince a real person and create a trust connection necessary to conclude a job interview, offering a deeper analysis for the last step of this attack life cycle validation.

References

Al-Charchafchi, A., Manickam, S., and Alqattan, Z. N. (2019). Threats against information privacy and security in social networks: A review. In *International Conference on Advances in Cyber Security*, pages 358–372. Springer.

Aroyo, A. M., Rea, F., Sandini, G., and Sciutti, A. (2018). Trust and social engineering in human robot interaction. *IEEE Robotics and Automation Letters*, 3(4):3701–3708.

Assenmacher, D., Clever, L., Frischlich, L., Quandt, T., Trautmann, H., and Grimme, C. (2020). Demystifying social bots: On the intelligence of automated social media actors. *Social Media+ Society*, 6(3):2056305120939264.

Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. (2011). The socialbot network: when bots socialize for fame and money. In *Proceedings of the 27th annual computer security applications conference*, pages 93–102.

Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. (2013). Design and analysis of a social botnet. *Computer Networks*, 57(2):556–578.

Camisani-Calzolari, M. (2012). Analysis of twitter followers of the us presidential election candidates: Barack obama and mitt romney. *http://digitalevaluations. com*.

Castells, M. (2009). Communication power. nueva york: oxford university press.

Crossler, R. and Bélanger, F. (2014). An extended perspective on individual security behaviors. *ACM SIGMIS Database*, 45(4):51–71.

Culot, G., Fattori, F., Podrecca, M., and Sartor, M. (2019). Addressing industry 4.0 cybersecurity challenges. *IEEE Engineering Management Review*, 47(3):79–86.

Darwish, A., Zarka, A. E., and Aloul, F. (2012). Towards understanding phishing victims' profile. In *2012 International Conference on Computer Systems and Industrial Informatics*, pages 1–5.

De Kimpe, L., Ponnet, K., Walrave, M., Snaphaan, T., Pauwels, L., and Hardyns, W. (2020). Help, i need somebody: Examining the antecedents of social support seeking among cybercrime victims. *Computers in Human Behavior*, 108:106310.

Dewangan, M. and Kaushal, R. (2016). Socialbot: Behavioral analysis and detection. In *International Symposium on Security in Computing and Communication*, pages 450–460. Springer.

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7):96–104.

Freitas, C., Benevenuto, F., Ghosh, S., and Veloso, A. (2015). Reverse engineering socialbot infiltration strategies in twitter. In *IEEE/ACM ASONAM 2015)*, pages 25–32. IEEE.

Freitas, C., Benevenuto, F., and Veloso, A. (2014). Socialbots: Implicações na segurança e na credibilidade de serviços baseados no twitter. *SBRC, Santa Catarina, Brasil*, pages 603–616.

Gallegos-Segovia, P. L., Bravo-Torres, J. F., Larios-Rosillo, V. M., Vintimilla-Tapia, P. E., Yuquilima-Albarado, I. F., and Jara-Saltos, J. D. (2017). Social engineering as an attack vector for ransomware. In *CHILECON 2017*, pages 1–6. IEEE.

Greitzer, F. L., Purl, J., Leong, Y. M., and Sticha, P. J. (2019). Positioning your organization to respond to insider threats. *IEEE Engineering Management Review*, 47(2):75–83.

Grimme, C., Preuss, M., Adam, L., and Trautmann, H. (2017). Social bots: Human-like by means of human control? *Big data*, 5(4):279–293.

Guzman, A. L. and Lewis, S. C. (2020). Artificial intelligence and communication. *New Media & Society*, 22(1):70–86.

Hepp, A. (2020). Artificial companions, social bots and work bots. *Media, Culture & Society*, 42(7-8):1410–1426.

Huber, M., Kowalski, S., Nohlberg, M., and Tjoa, S. (2009). Towards automating social engineering using social networking sites. In *2009 International Conference on Computational Science and Engineering*, volume 3, pages 117–124. IEEE.

Khan, R. and Das, A. (2018). Build better chatbots. *A complete guide to getting started with chatbots*.

Klimburg-Witjes, N. and Wentland, A. (2021). Hacking humans? social engineering and the construction of the "deficient user" in cybersecurity discourses. *Science, Technology, & Human Values*, 46(6):1316–1339.

Messias, J., Benevenuto, F., and Oliveira, R. (2018). Bots sociais: Como robôs podem se tornar pessoas influentes no twitter? *Revista Eletrônica de Iniciação Científica em Computação*, 16(1).

Mitnick, K. D. and Simon, W. L. (2003). *The art of deception*. John Wiley & Sons.

Paradise, A., Shabtai, A., and Puzis, R. (2019). Detecting organization-targeted socialbots by monitoring social network profiles. *Networks and Spatial Economics*, 19(3):731–761.

Piovesan, L. G., Silva, E. R. C., de Sousa, J. F., and Turibus, S. N. (2019). Engenharia social: Uma abordagem sobre phishing. *REVISTA CIENTÍFICA DA FACULDADE DE BALSAS*, 10(1):45–59.

Rouse, M. (2013). What is socialbot? *WhatIs.com*.

Salahdine, F. and Kaabouch, N. (2019). Social engineering attacks: a survey. *Future Internet*, 11(4):89.

Shafahi, M., Kempers, L., and Afsarmanesh, H. (2016). Phishing through social bots on twitter. In *2016 IEEE International Conference on Big Data*, pages 3703–3712. IEEE.

Stoeckli, E., Uebernickel, F., and Brenner, W. (2018). Exploring affordances of slack integrations and their actualization within enterprises-towards an understanding of how chatbots create value. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Tiwari, V. (2017). Analysis and detection of fake profile over social network. In *ICCCA 2017*, pages 175–179. IEEE.

Ye, W. and Li, Q. (2020). Chatbot security and privacy in the age of personal assistants. In *IEEE/ACM SEC 2020*, pages 388–393. IEEE.