Ataques Automatizados de Engenharia Social com o uso de *Bots* em Redes Sociais Profissionais

Maurício Ariza[1], **Antônio João G. de Azambuja**[1],
Jéferson C. Nobre[1], **Lisandro Z. Granville**[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – RS – Brazil

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

Abstract. *Virtual human interactions have been intensified with the increasing use of the Internet and social networks, raising the risk of Social Engineering cyber threats. The usage of Bots in those attacks allow scalability in the exploitation of users trust, causing security risks. There are few papers focusing on automated Social Engineering actions using Bots. This paper presents an assessment of the controls used in a professional social network to identify and block automated attacks, using a Bot as a proof of concept. The analysis and discussion of the results allow demonstrating the security vulnerabilities present in professional networks that can be exploited to build a trust relationship between the user and a malicious Bot.*

Resumo. *As interações humanas virtuais têm sido ampliadas com o uso crescente da Internet e redes sociais, elevando os riscos de ameaças cibernéticas de Engenharia Social. O uso de Bots nesses ataques permite escalabilidade na exploração da confiança dos usuários, provocando riscos de segurança. Poucos são os trabalhos com foco nas ações automatizadas de Engenharia Social com o uso de Bots. Este artigo apresenta uma verificação dos controles de uma rede social profissional quanto à identificação e bloqueio desses ataques automatizados, utilizando um Bot de prova de conceito. A análise e discussão dos resultados permite demonstrar as vulnerabilidades de segurança presentes nas redes profissionais que podem ser exploradas para construção da relação de confiança do usuário com um Bot malicioso.*

1. Introdução

Ataques cibernéticos exploram as vulnerabilidades das estruturas de Tecnologia da Informação e Comunicação (TIC), incluindo as redes sociais, que se estabeleceram ao longo dos anos como ferramentas de interação humana [Shires 2018]. No entanto, essas redes sociais emergem desafios e riscos relacionados à Segurança Cibernética (SegCiber) [Klimburg-Witjes and Wentland 2021].

As redes sociais fornecem serviços *on-line*, coletam dados pessoais e corporativos formando uma base de dados de alto valor, sendo passível de ser utilizada como ferramentas para ataques cibernéticos que exploram as relações de confiança [Crossler and Bélanger 2014].

Essas relações de confiança no ambiente cibernético têm proporcionado um cenário para a prática de atos ilícitos, implicando em riscos de SegCiber. Os ataques têm explorado a interação humana em conjunto com as brechas

tecnológicas, enfraquecendo a cadeia de segurança [Salahdine and Kaabouch 2019] [Klimburg-Witjes and Wentland 2021]. Na prática, o fator humano é o elo mais fraco na cadeia de SegCiber [Mitnick and Simon 2003]. A interconectividade das redes sociais e o crescimento da dimensão cognitiva do trabalho estão tornando os recursos humanos como um dos pilares da segurança [Culot et al. 2019] [Greitzer et al. 2019].

As organizações têm empregado soluções de defesa para enfrentar os ataques cibernéticos, tais como *Firewalls*, Sistema de Detecção de Intrusão (*Intrusion Detection System* - IDS), Sistema de Prevenção a Intrusão (*Intrusion Prevention System* - IPS) e Antivírus. No entanto, esses mecanismos de defesa não têm sido suficientes para impedir integralmente as ações de Engenharia Social (ES) no ambiente cibernético. Os atacantes que utilizam ES vêm adotando mecanismos automatizados para explorar as relações de confiança, tendo como objetivo obter dados e informações relevantes de potenciais alvos. Os ataques automatizados requerem pouca intervenção humana e são desenvolvidos visando simular o comportamento humano [Huber et al. 2009] [Shafahi et al. 2016].

O crescente aumento do uso das redes sociais para estabelecer relacionamentos pessoais e profissionais abre um campo para as ações de *Bots* de Engenharia Social Automatizada (ESA) [Huber et al. 2009]. Os *Bots* são *softwares* automatizados, que por vezes utilizam recursos como inteligência artificial (IA), e são capazes de executar comandos de operação e controle, sem a necessidade de participação humana. São ferramentas com a capacidade de se passar por seres humanos, imitando as atividades dos usuários reais [Shafahi et al. 2016].

Bots podem ser utilizados para ações positivas, como por exemplo, ajudar o usuário na sua experiência *on-line*. Para os autores [Dickerson et al. 2014] os humanos tendem a confiar nos *Bots*. Contudo, os *Bots* de ESA têm sido utilizados como uma ferramenta para ataques de ES, já que são escaláveis, permitindo que um único atacante contate um grande número de potenciais vítimas simultaneamente, na busca de informações confidenciais [Huber et al. 2009][Dewangan and Kaushal 2016].

Na literatura, poucos trabalhos apresentam análises sobre a ESA com o uso de *Bots*. A maioria dos trabalhos estuda a área da psicologia social, com foco no comportamento humano diante das ações de ES [Huber et al. 2009]. Os autores [Al-Charchafchi et al. 2019] e [Piovesan et al. 2019] abordam as ameaças à segurança nas redes sociais, decorrentes dos ataques de ES utilizando contas falsas, roubo de identidade e *phishing*. No sentido de influenciar os usuários nas redes sociais, há trabalhos que avaliam as vulnerabilidades das redes sociais com o uso de *SocialBots* para campanhas de convencimento nas redes. Os autores [Freitas et al. 2014] e [Messias et al. 2018] analisam o uso de *Bots* no *Twitter* para influenciar os usuários e comprometer a estrutura da rede. Já [Huber et al. 2009] propõem a automação das tarefas de ES por meio de um *Bot* no *Facebook*, concluindo que a persuasão é um recurso essencial no processo de ESA.

Nesse contexto onde o uso de *Bots* de ESA permite automação e escalabilidade dos ataques com menor exposição do agente malicioso em si, este trabalho se propõe a analisar a mais popular rede social profissional atualmente, o *LinkedIn*, e verificar se a mesma oferece controles que possam impedir ou dificultar a ação automatizada de ataques de Engenharia Social. As principais contribuições desse trabalho são: i) avaliar a capacidade de detecção e bloqueio de ataques automatizados por parte da rede social *LinkedIn*;

ii) implementar uma prova de conceito para validar a viabilidade técnica desses ataques; e iii) propor melhorias que possam ser utilizadas por essas redes a fim de diminuir os riscos de Engenharia Social Automatizada aos seus usuários.

O artigo inicialmente aborda, na Seção 2, os conceitos relacionados com a teoria para embasar a pesquisa. Na Seção 3, analisa os trabalhos relacionados com o tema da pesquisa. A seguir, na Seção 4, discorre sobre a apresentação do problema e o método de ataque. Na Seção 5, apresenta o protótipo, experimento e discussão dos resultados. A seguir, na Seção 6 as limitações da pesquisa são mencionadas. Por fim, apresenta-se na Seção 7 a conclusão e uma abordagem para trabalhos futuros.

2. Referencial Teórico

Embora sejam limitados os trabalhos no tema do uso de ataques automatizados de Engenharia Social com o uso de *Bots*, existem publicações relacionadas ao tema que ajudam a embasar teoricamente as premissas utilizadas para condução do trabalho e proposta.

2.1. Engenharia Social Automatizada

A ES refere-se à exploração do comportamento humano no tocante ao uso dos sistemas de informações para obtenção de dados e informações relevantes de potenciais alvos. Os ataques de ES colocam o atacante em uma posição favorecida no fluxo de informações, tirando proveito de uma relação de confiança [Mitnick and Simon 2003]. O desenvolvimento de uma relação de confiança faz uso da manipulação psicológica induzindo as pessoas realizarem ações específicas. Conquistar a confiança das vítimas é um objetivo dos engenheiros sociais.

Esses ataques podem ser detectados, no entanto, não são facilmente interrompidos [Libicki 2018]. As técnicas de ES direcionam os usuários a responderem solicitações sem uma análise adequada as informações disponibilizadas, seguindo 4 (quatro) estágios, a saber: i) obter as informações sobre a vítima para uma primeira abordagem; ii) estabelecer uma relação de confiança entre o atacante e a vítima; iii) explorar as informações para o desenvolvimento de ações específicas; e iv) executar o ataque para alcançar o(s) seu(s) objetivo(s) [Mitnick and Simon 2003] [Tioh et al. 2019].

Os ataques de ES demandam tempo e recursos para estabelecer um relacionamento de confiança. No entanto, o desenvolvimento de uma *interface* homem-máquina permite que tais relacionamentos sejam automatizados [Guzman and Lewis 2020]. Os engenheiros sociais podem utilizar a automação para desenvolver ferramentas pré-programadas para realizar tarefas sem a intervenção humana, possibilitando a escalabilidade dos ataques de ES [Huber et al. 2009] [Shafahi et al. 2016].

Os ataques automatizados podem ser preparados utilizando informações de valor e/ou influenciando determinados grupos nas redes sociais [Mitnick and Simon 2003] [Gallegos-Segovia et al. 2017]. Essas redes representam um espaço virtual atrativo para os atacantes explorarem as vulnerabilidades técnicas e a falta de conhecimento e conscientização dos usuários sobre ações de ES [Al-Charchafchi et al. 2019]. Uma das vulnerabilidades que são encontradas em redes sociais é a criação de perfis falsos, os quais constituem um percentual significativo dos usuários dessas redes [Tiwari 2017]. O Relatório de Investigação de Violação de Dados, publicado em 2021, descreve que 40 %

dos casos de violação dos dados tem relação com as ações de ES [1].

2.2. *Bots*

Bot é o termo resumido da palavra da língua inglesa *Robot*, que na tradução livre significa Robô. É uma ferramenta automatizada que realiza uma série de funções pré-programadas de operação e controle. Os *Bots* podem ser autênticos, que têm como objetivo realizar atividades úteis para os usuários, por outro lado também existem *Bots* de cunho malicioso, que podem realizar ataques para obter informações relevantes ou manter o controle do dispositivo acessado. *Bots* podem ser utilizados para ações de disseminação de informações falsas (*fake news*), *spam* e *phishing* [Freitas et al. 2015].

No contexto de ESA os cibercriminosos usam os *Bots* maliciosos para simular o comportamento humano, burlando os mecanismos de segurança. Com o crescimento das redes sociais e o grande volume de dados no ciberespaço, os engenheiros sociais passaram a espalhar *Bots* com comportamento semelhante ao do ser humano para um grande número de usuários. Esses *Bots* simulam conversas humanas, conhecidos como *ChatBots* e, os que atuam nas redes sociais, os *SocialBots* [Shafahi et al. 2016].

ChatBot é a integração de sistemas, ferramentas e roteiros que promovem conversas por mensagens instantâneas com ou sem a participação de humanos [Stoeckli et al. 2018]. São desenvolvidos para ajudar usuários humanos em situações de serviços específicos, não sendo exaustivo. Por exemplo: atendimento ao cliente, atendimento por telefone e serviço de educação digital [Grimme et al. 2017]. O uso da linguagem natural nos *ChatBots* é um desafio a ser superado para o desenvolvimento dessa ferramenta [Khan and Das 2018].

SocialBot é uma ferramenta de *software* que simula o comportamento humano para realizar interações automatizadas nas redes sociais [Rouse 2013]. Os *SocialBots* têm a capacidade de comprometer a estrutura das redes sociais, influenciando os usuários e aumentando o número de seguidores, para inflar os índices de popularidade de uma determinada conta de perfil [Camisani-Calzolari 2012]. Essa ferramenta é eficaz para ataques de ES, utilizando-se de informações sensíveis de possíveis vítimas, como o roubo de identidade [Dewangan and Kaushal 2016].

Essas ferramentas têm sido desenvolvidas com a ajuda de mecanismos de IA que interagem com os usuários [Freitas et al. 2015]. A IA é similar a inteligência humana, desenvolvida com a automatização conforme a necessidade da aplicação [Ferrara et al. 2016]. Na medida que um certo grau de inteligência é incorporado nas ferramentas para simular o comportamento humano, aumenta a capacidade e escalabilidade dos ataques.

3. Trabalhos relacionados

A proposta de desenvolvimento de um *Bot* automatizado para ataques de ES demanda uma análise estruturada da literatura para apoiar este estudo. A análise permitiu identificar que parte dos estudos de ES têm foco no comportamento humano diante das ações de *SocialBots*, *phishing* e *spam*.

[1] https://enterprise.verizon.com/resources/reports/2021/2021-data-breach-investigations-report.pdf

Os autores [Dewangan and Kaushal 2016] abordam o uso de *SocialBots* em campanhas políticas e *marketing* de produtos. No trabalho os autores apontam os riscos de segurança atrelados a essa prática, no que diz respeito ao acesso às informações pessoais dos usuários. Diante dos riscos o trabalho menciona a necessidade de identificação desses *SocialBots*. Esse procedimento permite assegurar a reputação de uma rede social que está sendo objeto do ataque. Para tanto, os autores desenvolveram um modelo de detecção desses *SocialBots* considerando a análise de comportamento.

O trabalho dos autores [Aroyo et al. 2018], discorre como a ES explora a relação de confiança entre os usuários e *Bots*. Com base no modelo de [Mitnick and Simon 2003] foi desenvolvido um *Bot* para simular um ataque de ES. Inicialmente o *Bot* buscou obter informações com perguntas de cunho privado. Na sequência estabeleceu uma relação de confiança com os participantes, por meio dos *Bots*, para uma aproximação anônima com o alvo. Nos resultados do estudo os participantes, na sua maioria (62%), demonstraram confiança na ferramenta mencionando o comportamento ético, tendo em vista que foi desenvolvida considerando questões éticas.

O artigo *Threats Against Information Privacy and Security in Social Networks: A Review* [Al-Charchafchi et al. 2019], apresenta uma revisão das pesquisas sobre privacidade e ameaças à segurança nas redes sociais. Para os autores, no que pese a literatura apresentar trabalhos sobre privacidade, mais esforços são necessários. O ambiente das redes sociais é uma rica fonte de dados pessoais, tornando-se um atrativo para ações dos engenheiros sociais, que exploram a falta de conscientização e conhecimento dos usuários nas questões relacionadas com a segurança.

Para [Al-Charchafchi et al. 2019] a complexidade dos ataques de ES está alinhada com a possibilidade da combinação das estratégias sociais e técnicas para realizar um crime cibernético. Nessa linha os autores [Piovesan et al. 2019] afirmam que políticas de segurança podem oferecer maior nível da segurança da informação, no entanto não garantem proteção completa.

Os trabalhos dos autores [Freitas et al. 2014], [Messias et al. 2018] e [Shafahi et al. 2016], discutem o impacto do uso dos *SocialBots* no *Twitter* para caracterizar o comportamento da ferramenta em uma grande base de dados, medir a capacidade da ferramenta influenciar os usuários na rede, detectar automaticamente esses *Bots* e analisar os riscos de segurança decorrentes do uso de *phishing*, por meio de *SocialBots* no *Twitter*.

Nos resultados os autores [Freitas et al. 2014], destacam que o método por eles desenvolvido para caracterizar e detectar os *SocialBots*, teve um indicador de detecção com 92% de sucesso. Os autores [Messias et al. 2018], afirmam que um simples *Bot* pode alcançar altos níveis de influência no *Twitter*. Já [Shafahi et al. 2016], apontam a necessidade de aumentar o nível de conscientização sobre as ações de *phishing* que utilizam *SocialBots*. Os autores afirmam que essas ações constituem uma ameaça para as organizações.

Por fim, o trabalho dos autores [Huber et al. 2009], apresenta o ciclo de um ataque de ESA, com o uso de um *Bot*. O ataque demonstrou como as redes sociais podem ser utilizadas pelos engenheiros sociais para obter informações. Para tanto, no trabalho foram realizados 2 (dois) experimentos. O primeiro analisou a capacidade do *Bot* em

obter informações nas redes sociais. O segundo realizou o *Turing Test* [Turing 2009], que busca avaliar a capacidade de uma máquina imitar um ser humano.

Para os autores a ESA com *Bots* é escalável e requer menos recursos humanos. A ferramenta foi utilizada em uma prova de conceito no *Facebook*. Os 2 (dois) experimentos permitiram ratificar que é possível automatizar ações de ES para obter informações e demonstrar que o *Bot* utilizado não foi identificado pelas medidas de segurança do *Facebook*. O número crescente das interações sociais dos usuários nas redes, torna os *Bots* de automação de ES uma ferramenta interessante para os engenheiros sociais.

4. Solução Proposta

Ataques de ES em redes sociais já são conhecidos e documentados, pois são espaços de interação cujas características despertam grande interesse para agentes maliciosos. Em especial, a capacidade de personificar com facilidade algum personagem que possa ganhar a confiança da vítima [Crossler and Bélanger 2014].

Diferentemente de outras redes sociais como *Facebook*, *Twitter* e *Instagram*, as redes sociais profissionais promovem uma atmosfera de ambiente corporativo, focada em conexões e relacionamentos para crescimento na carreira. Essas redes despertam o interesse de recrutadores e empresas na busca por candidatos para suas vagas e perfis de clientes em potencial. Nesse contexto, elas apresentam um cenário que inspira maior credibilidade e confiança entre os seus usuários, tornando esse grupo alvos em potencial para ataques direcionados e complexos.

Embora não referenciada em trabalhos acadêmicos, uma forma de ataque de ES existente nessas redes se caracteriza pela criação de um perfil falso na rede profissional por um atacante, entrando então em contato com potenciais vítimas se apresentando como recrutador para uma oportunidade de trabalho[2]. A partir do interesse da vítima, o atacante realiza entrevistas no intuito de roubar informações. Algumas formas comuns são descobrir informações confidenciais de empresas ou projetos onde a vítima tenha trabalhado, ou, ao final do processo, oferecer um falso contrato de trabalho, solicitando dados pessoais e uma cópia do passaporte ou documento similar, informações que podem ser utilizadas para roubo de identidade. Todo o processo envolvido e o contato com a vítima são feitos de forma manual pelo atacante.

A ES tem sua questão central no campo da psicologia, tendo a computação como uma ferramenta para viabilizar o trabalho do atacante. O nosso trabalho busca implementar uma prova de conceito para validar a viabilidade técnica de ataques de ESA pela ausência ou insuficiência de controles de segurança nas redes sociais profissionais, fato que abre brechas para o trabalho de agentes maliciosos.

A Seção 8.2 da Política de Uso do *LinkedIn* [3] especifica quais ações são permitidas ou proibidas na plataforma, como o uso de informações falsas no perfil ou o uso de *Bots* e *Scripts*. Porém, parte dessas regras são aplicadas apenas através de denúncias por parte de outros usuários, e não por controles técnicos.

A ausência de formas de controle ou validação permite um usuário identificar-se

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

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

como funcionário de qualquer empresa, mencionar habilidades em diferentes campos de estudo ou construir um perfil que possa ser do seu interesse. Embora proibidos, códigos de automatização são amplamente utilizados, sendo possível encontrá-los em repositórios públicos como o *GitHub*. Levando em conta apenas essas duas questões, um atacante pode: (i) criar um perfil que atraia o interesse de seus alvos; e (ii) utilizar técnicas de automatização para aumentar a escalabilidade do seu ataque, sem a necessidade de burlar os mecanismos de controle da plataforma.

4.1. Método de Ataque

Tendo como referência o trabalho feito por [Huber et al. 2009], buscamos na pesquisa validar se os mecanismos e processos de controle da rede social *LinkedIn* são capazes de identificar e bloquear um ataque de ES automatizado, onde as ações do fluxo de ataque são realizados utilizando um *Bot* visando diversos alvos simultaneamente, sem a interação manual direta entre o atacante e a vítima.

O método para testar a proposta utiliza como base o ataque apresentado no início da Seção 4, onde o atacante apresenta-se como um recrutador. Para isso, foi criado um perfil falso e um *Bot* com o objetivo de: i) realizar a busca de perfis de potenciais vítimas; ii) adicionar de forma automatizada uma grande quantidade de vítimas como contatos; e iii) enviar de forma simultânea mensagens às vítimas oferecendo uma falsa oportunidade de trabalho como chamariz.

O *Bot* precisa, portanto, ser capaz de identificar um grande número de usuários simultaneamente a partir de palavras-chave de interesse do atacante e entrar em contato com todos eles, sem ser identificado e bloqueado pelos controles da rede social. Para tanto, é necessário uma infraestrutura técnica, com a combinação de uma plataforma de rede social e requisitos para automação do comportamento de uma conta, utilizando uma *Application Programming Interface* (API - Interface de Programação de Aplicativo, tradução livre) ou mecanismos proprietários para interagir com a plataforma [Assenmacher et al. 2020].

5. Avaliação

5.1. Protótipo

O *LinkedIn* oferece uma API bastante completa para interação com a plataforma, sendo, portanto o caminho natural para uma interação feita através de *software*. Por decisão dos pesquisadores optou-se por replicar o comportamento de um usuário padrão via navegador realizando diversas ações simultâneas, caracterizando claramente o uso de automatização e violação das políticas de uso da rede social.

Para verificar a viabilidade do ataque, os autores então desenvolveram uma aplicação de prova de conceito em linguagem *Python* para interagir com a rede social, utilizando a biblioteca *Selenium* a fim de realizar as requisições diretamente através de um navegador.[4].

Foi criado um perfil falso do atacante utilizando informações aleatórias, valendo-se da ausência de validação das informações pela rede social. Dados como graduações e

[4]Por questões éticas na publicação de uma ferramenta de ataque, os autores optaram por manter o código da mesma em um perfil privado do GitHub, podendo fornecer acesso sob requisição.

diplomas, nível de conhecimento e experiências profissionais atuais e prévias, associando o indivíduo diretamente à empresas e instituições verdadeiras, podem ser registrados pelo atacante sem dificuldade. Toda essa construção auxilia na demonstração da veracidade permitindo passar confiança às vítimas.

O protótipo considerou um modelo para os estágios do ataque de ES [Mitnick and Simon 2003], sendo estruturado em 3 (três) etapas: i) Autenticação: para o acesso à rede social; ii) Busca: para mapeamento dos alvos; e iii) Abordagem: para contatar as vítimas. As etapas são detalhadas nos parágrafos a seguir.

1ª. Etapa de Autenticação: Essa fase busca verificar se a rede social é capaz de detectar o processo de autenticação de um usuário sendo executado de forma automatizada. Para isso, o *Bot* desenvolvido inicia o navegador e abre a página do *LinkedIn*, sendo automaticamente direcionado à página de *login*. O código-fonte da página é então mapeado para que os campos de autenticação sejam identificados. São solicitados o usuário e a senha do perfil do atacante, os quais são introduzidos diretamente nos campos apropriados a fim de acessar o página principal de um usuário autenticado.

Figura 1. Fluxo de ataque - Etapa de Autenticação

2ª. Etapa de Busca: Essa fase busca verificar se a rede social é capaz de detectar a realização de diversas buscas realizadas simultaneamente de forma automatizada. Para isso, o *Bot*, já com o usuário do atacante autenticado, recebe palavra(s)-chave de busca. Novamente o código-fonte da página é mapeado, o campo de busca na *interface* do usuário é identificado e os termos são introduzidos no mesmo, fazendo com que o *LinkedIn* retorne uma lista de perfis baseada naqueles critérios. Embora exista uma relação entre o termo buscado e os resultados, a quantidade e ordem dos perfis exibidos como resultados da busca são definidos unicamente pelo algoritmo da própria rede social. Portanto para os objetivos desse trabalho os resultados da pesquisa em si são armazenados meramente para criação de um banco de dados de alvos pelo atacante, não fazendo parte desse escopo analisar o algoritmo da rede social e seu comportamento quanto ao retorno de resultados.

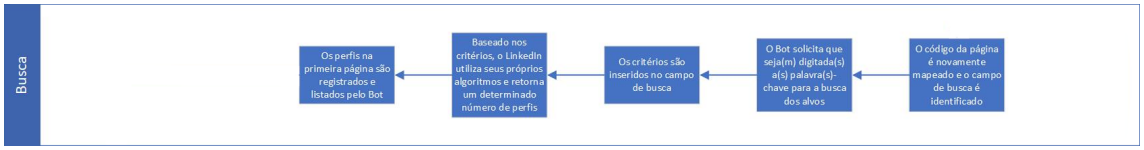

Figura 2. Fluxo de ataque - Etapa de Busca

3ª. Etapa de Abordagem: Essa fase busca verificar se a rede social é capaz de detectar o envio simultâneo de mensagens a diferentes usuários de forma automatizada. Para isso, o *Bot* recebe uma mensagem personalizada que será enviada para as vítimas. Mapeando o código-fonte da página onde se apresentam os resultados da pesquisa feita

na etapa anterior, para cada uma das vítimas é identificado e acionado o botão para solicitar conexão com a mesma, a mensagem personalizada é introduzida no conteúdo da solicitação, utilizando parâmetros customizados para garantir que cada mensagem chame a vítima pelo próprio nome. Esse procedimento é repetido para todos os demais perfis listados.

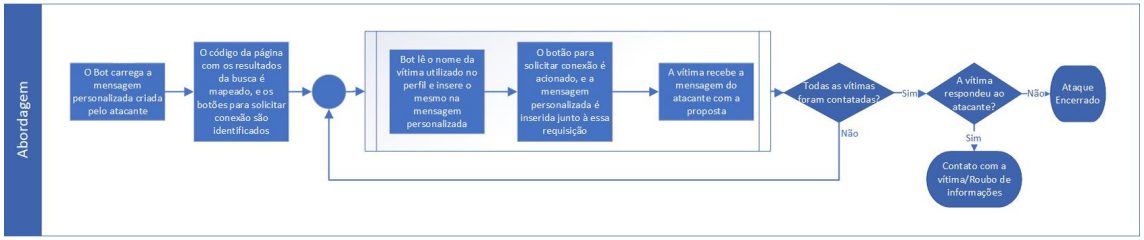

Figura 3. Fluxo de ataque - Etapa de Abordagem

O uso de palavras-chave que identifiquem determinados tipos de profissionais - como um cargo ou habilidade técnica específica - pode ser uma forma de refinar os resultados da busca e definir um tipo de perfil específico de alvos. Porém, do ponto de vista do ataque a determinação de termos exatos a serem utilizados é de menor relevância, visto que o objetivo do atacante é mapear o maior número possível de vítimas.

5.2. Experimentos

Os testes foram realizados de acordo com as etapas de autenticação, busca e abordagem, detalhadas no Fluxo de Ataque. O código foi testado em uma máquina *Windows*, utilizando a versão 3.9 do *Python*, a versão 3.141.0 da biblioteca *Selenium* com o *driver* na versão 90.0.4430.24 do *driver* para o navegador *Google Chrome*, que por sua vez estava na versão 90.0.4430.72.

Testes da 1ª. Etapa - Autenticação: Sendo o objetivo dessa etapa verificar a capacidade da rede social identificar a autenticação de forma automatizada, configuramos o *Bot* de testes com as credenciais da conta criada para o atacante. O mesmo foi capaz de realizar a autenticação com sucesso mesmo com o processo sendo repetido diversas vezes em sequência ou de forma simultânea sem exibição de mensagens de erro, solicitação de controles do tipo *Captcha* ou quaisquer indícios que a rede tenha detectado aquele acesso como tendo sido realizado de forma automatizada, reforçando que pelo uso da biblioteca *Selenium* todas as ações são realizadas diretamente através do navegador. Para fins comparativos, os testes foram repetidos com variações como credenciais inseridas manualmente na execução, credenciais lidas em arquivo e uso de credenciais inválidas, não tendo sido observados diferenças de resultado além de sucesso quanto utilizadas credenciais válidas e erro quando utilizadas inválidas.

Testes da 2ª. Etapa - Busca: Para validar a resposta da rede social quanto à execução de buscas de forma automatizada, o *Bot* foi alimentado com diferentes palavras-chave, sendo executadas buscas de forma contínua e simultânea a fim de verificar alterações de comportamento por parte da mesma que pudessem indicar detecção de comportamento automatizado, não tendo sido observado nenhuma ação por parte dos pesquisadores. Não foram executados testes de estresse/carga pois entende-se que o objetivo de um atacante seja atingir o maior número possível de vítimas sem ser identificado, e não de causar negação de serviço na aplicação.

Foram realizados em menor escala alguns testes quanto à precisão do resultados, utilizando os termos ”teste”, ”Engenharia Social”, ”Bot”, ”Redes Sociais”, ”Segurança da Informação”e ”Python”, verificando uma amostra dos perfis retornados a fim de verificar a relevância e as diferenças quanto à quantidade de perfis retornados. É importante salientar porém que o objetivo dos pesquisadores foi identificar potencial mudança de comportamento por conta das buscas estarem sendo executadas através de um *Bot*, pois os resultados das buscas em si são resultado do próprio algoritmo do *LinkedIn*, não tendo influência direta por parte dos pesquisadores além do termo de pesquisa utilizado e sendo fora do escopo deste trabalho testar a capacidade de resposta do mesmo.

Testes da 3ª. Etapa - Abordagem: A terceira e última etapa seria verificar se ocorreria a detecção com a abordagem de indivíduos. Esta etapa era a mais diretamente afetada pelas questões éticas envolvidas nesse trabalho, que são discutidas em mais detalhes na Seção 6. Para execução dos testes, o *Bot* capturava os perfis retornados nas buscas da etapa anterior - a fim de limitar o número de alvos o código foi configurado para filtrar apenas os resultados presentes na primeira página de busca, sendo entre 15 e 21 perfis por palavra-chave. A cada rodada todos os alvos eram contatados simultaneamente, recebendo uma solicitação de conexão e uma mensagem personalizada, tendo sido verificado também a execução de duas rodadas simultaneamente. Em nenhum momento foi identificado novamente ações por parte da rede social. Após o envio, os pedidos de conexão e as mensagens eram automaticamente cancelados e excluídos antes da interação com os alvos ocorrer.

5.3. Discussão

Uma das contribuições deste trabalho foi a busca pela validação da hipótese de ausência de controles por parte das redes sociais, que embora seja conhecida no meio da tecnologia, não encontramos referências na academia.

Como mencionado na Seção 4, não é um dos objetivos desta pesquisa analisar as vulnerabilidades dos usuários em si e os aspectos psicológicos explorados pela Engenharia Social, e sim como a ausência ou ineficiência dos controles utilizados pelas redes facilitam os ataques e fornecem oportunidade de escalabilidade. Sendo assim foi possível realizar a prova de conceito com a aplicação do *Bot*, demonstrando o potencial de uso real para uma eventual atividade maliciosa.

No caso das redes sociais em geral, um dos principais argumentos contra o uso de controles mais rígidos é o impacto que os mesmos terão na usabilidade, podendo levar os usuários a migrarem para plataformas concorrentes. Porém é possível encontrar um balanço entre as necessidades, trazendo maior segurança aos usuários com um mínimo impacto.

A possibilidade de ataques de Engenharia Social Automatizada é um exemplo. Como já dito, atualmente a identificação e remoção desses usuários ocorre apenas a partir de denúncias. Partindo do princípio que a própria Política de Uso proíbe o uso de automação, quaisquer comportamentos que indicassem essas características poderiam ser bloqueados, por exemplo:

- Mais de um *login* simultâneo do usuário, ou diversos *logins* com sucesso seguidos;
- Quantidade de requisições simultâneas e contínuas acima da capacidade de serem produzidas por um ser humano utilizando a plataforma;e

- Adicionar como contatos ou enviar mensagens para grandes quantidades de usuários simultaneamente ou em uma janela curta de tempo (que também poderia indicar *SPAM*).

Considerando que certos serviços necessitam utilizar algumas ferramentas de automação - como recrutadores reais - esses controles poderiam ser mais rígidos nas conexões via navegador (onde o uso esperado é de ser feito por uma pessoa) e mais flexíveis via API (onde é possível inclusive ter um melhor monitoramento por parte da plataforma). Esse formato permitiria oferecer um determinado número de requisições sem custo para usuários menores e planos mais robustos com a contratação de serviços profissionais da plataforma, como o já existente *LinkedIn Recruiter*.

Controles que bloqueiem automação, como os sugeridos, terão pouco ou nenhum impacto no uso de usuários regulares. Além de diminuírem grandemente os riscos de ataques maliciosos automatizados, também reduziriam o número de *SPAMs* e demais serviços não solicitados que, embora violem as Políticas de Uso, ocorrem diariamente na plataforma.

Uma questão mais complexa porém é a facilidade de criação de perfis falsos, problema enfrentado pelas redes sociais no geral. Com a cultura de expansão das suas redes de contatos, não é difícil que um perfil novo tenha rapidamente conexões suficientes para demonstrar credibilidade, sem contar a possibilidade da criação de diversos perfis falsos que gerem credibilidade uns aos outros através de depoimentos e recomendações.

Mas como gerar essa credibilidade sem processos de validação complexos?

Levando em conta que discussões sobre obrigar a identificação dos usuários já esteja ocorrendo em outras redes como o *Twitter*[5], onde podemos dizer que as características de uso são um tanto diferentes de redes profissionais, em uma rede cuja missão é conectar os profissionais do mundo para torná-los mais produtivos e bem-sucedidos [6], não seria ainda mais importante o interesse na credibilidade dos usuários? Caso não seja possível aplicar para todos os usuários, um bom começo seria exigir a validação de usuários que atuem como recrutadores na plataforma, oferecendo tanto uma forma de maior reconhecimento para esses profissionais quanto tornando mais difícil a personificação desses papéis.

6. Limitações

Após casos famosos como os experimentos de Milgran nos anos 70, estudos que envolvam o engano de pessoas enfrentam fortes dilemas éticos em sua produção. A exposição de pessoas reais a situações onde as mesmas serão iludidas, tendo suas vulnerabilidades exploradas sem seu consentimento, potencialmente podem gerar frustração e estresse psicológico após sua realização. Trabalhos no campo da ES, embora normalmente utilizem a tecnologia como suporte, implicam nessas mesmas questões de estudos psicológicos, sendo portanto necessário uma forte atenção dos autores e algumas limitações aos experimentos práticos.

Foram analisadas diversas possibilidades de execução de testes, sendo muito claras as diversas limitações em quaisquer delas para que os aspectos éticos fossem respei-

[5]https://www.cnnbrasil.com.br/business/elon-musk-diz-que-quer-todos-os-humanos-reais-verificados-no-twitter/

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

tados. Por conta disso, optou-se por focar individualmente em cada uma das etapas e testá-las separadamente levando em conta o ponto de vista da ausência de controles da plataforma. Assim seria possível validar as condições necessárias para que um ataque de Engenharia Social Automatizada ocorresse, sem a necessidade de realização de um ataque de ponta a ponta, onde seria necessário que as vítimas do teste acreditassem na história personificada a fim de que os resultados pudessem ser realmente validados.

Em especial a etapa de Abordagem foi onde ocorreram os maiores desafios, visto que a única forma de validar a ausência ou insuficiência de controles seria realizando a abordagem em si. Limitar os pedidos de conexão e mensagens a um número mínimo que fosse suficiente para caracterizar um comportamento automatizado, mas permitir um rápido controle de danos a fim de evitar o contato real com usuários, foi a forma encontrada para validar essa etapa em uma linha bastante tênue entre os limites éticos para uma pesquisa deste tipo.

Desta forma, mesmo levando em conta todas as limitações apresentadas, acreditamos ter sido possível validar as características necessárias para provar a viabilidade de um ataque automatizado de Engenharia Social.

7. Conclusão

Os ataques cibernéticos estão expondo as vulnerabilidades das redes computacionais. Os mecanismos de defesa não têm sido eficientes para impedir os ataques que exploram relações de confiança com o uso de *Bots*. O espaço virtual, no contexto das redes sociais, constitui-se um promissor cenário para a prática de toda sorte de atos ilícitos.

Este artigo explorou a ausência ou insuficiência de controles por parte dessas plataformas para detecção ou bloqueio dessas ameaças, apresentando como prova de conceito um *Bot* para simular ataques de ESA tendo como atrativo para os usuários ofertas de emprego.

As principais contribuições deste trabalho foram: i) implementar uma prova de conceito para validar a viabilidade técnica desses ataques de forma automatizada; e ii) apresentar e avaliar as descobertas do experimento dos ataques automatizados de ES.

O experimento demonstrou a viabilidade técnica para *Bots* de ESA, visto que foi possível realizar ações de forma remota e simultânea sem que houvesse qualquer restrição ou bloqueio por parte da plataforma. Os resultados apresentam o potencial de ferramentas similares para ações de ES, fato que demanda a necessidade de enfrentar os desafios impostos pelas questões de SegCiber.

Como trabalhos futuros a realização de provas de conceito com um *Bot* mais robusto identificaria os limites máximos de ações automatizadas suportados pela plataforma e uma emulação mais realista de um ataque automatizado de ponta a ponta. A implementação de capacidade de *chatbot* também verificaria a capacidade de interação, personificação e convencimento da vítima para fechamento do ciclo de um ataque de Engenharia Social.

Referências

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: Will a robot make you disclose sensitive information, conform to its recommendations or gamble? *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.

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

Crossler, R. and Bélanger, F. (2014). An extended perspective on individual security behaviors: Protection motivation theory and a unified security practices (usp) instrument. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 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.

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

Dickerson, J. P., Kagan, V., and Subrahmanian, V. (2014). Using sentiment to detect bots on twitter: Are humans more opinionated than bots? In *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pages 620–627. IEEE.

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 *2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 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 *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*, 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: A human–machine communication research agenda. *New Media & Society*, 22(1):70–86.

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*, page 0162243921992844.

Libicki, M. (2018). Could the issue of dprk hacking benefit from benign neglect? *Georgetown Journal of International Affairs*, 19:83–89.

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: Controlling the human element of security*. John Wiley & Sons.

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 (Big Data)*, pages 3703–3712. IEEE.

Shires, J. (2018). Enacting expertise: Ritual and risk in cybersecurity. *Politics and Governance*, 6(2):31–40.

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*.

Tioh, J.-N., Mina, M., and Jacobson, D. W. (2019). Cyber security social engineers an extensible teaching tool for social engineering education and awareness. In *2019 IEEE Frontiers in Education Conference (FIE)*, pages 1–5. IEEE.

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

Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the turing test*, pages 23–65. Springer.