Método para Testes Padronizados de Engenharia Social

Maurício Ariza[1], Afonso Comba de Araújo Neto[2]

[1]Universidade do Vale do Rio dos Sinos (UNISINOS)
Caixa Postal 275 – 93.022-000 – São Leopoldo – RS – Brazil

[2]CPD – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

mauariza@gmail.com,afonso@cpd.ufrgs.br

Abstract. *With more and more investments in security, the focus of the attackers has been the human factor. Social Engineering comprises of fraud and deception in order to gain information or to achieve the execution of actions of interest for the attacker, without the victim finding out that it is being manipulated. With the expansion of internet access, the number of frauds and victims are growing each day, highlighting the need for training of users and managers of computer systems. This study aims to test a methodology that offers a simple and intuitive way to perform tests for awareness and training of people for fraud protection.*

Resumo. *Com cada vez mais investimentos em segurança, o foco dos atacantes tem sido no fator humano. A Engenharia Social envolve basicamente a fraude e o engano para obter informações ou ações de interesse do atacante, sem que a vítima saiba que está sendo manipulada. Com a expansão do acesso à internet, o número de fraudes e vítimas tem crescido a cada dia, evidenciando a necessidade de capacitação dos usuários e gerentes de sistemas computacionais. Este trabalho objetiva testar um método que ofereça uma maneira simples e intuitiva de realizar testes para conscientização e capacitação de pessoas para proteção contra fraudes.*

1. Introdução

Conforme pesquisa da PwC (2015) com executivos de TI e Segurança, foi verificado aumento de 24% nos orçamentos de segurança e de 38% no número de incidentes detectados, porém foi também percebido um aumento de 56% no roubo de informações de propriedade intelectual altamente críticas, indicando possível crescimento em quantidade ou nível de especialização dos atacantes. A pesquisa ainda apresenta que embora o meio de invasão mais citado continue sendo os colaboradores, apenas 53% das empresas afirmam ter programas de treinamento e conscientização dos usuários [PWC 2015]. Com o aumento dos investimentos em segurança, é exigido de um potencial atacante cada vez mais conhecimento técnico e análise crítica do alvo para obter sucesso. Um atacante inteligente não desperdiçaria um longo tempo em uma atividade de alto grau de dificuldade se existissem formas alternativas de se atingir o mesmo objetivo em um tempo curto e com menor dificuldade [Hadnagy 2011]. Conforme Darwish et. al. (2012), o usuário, é considerado o "elo fraco da corrente", visto que a exploração de seus déficits tem sido a forma mais comum e bem-sucedida de

executar ataques. Conforme relatório da Verizon (2015), os casos de *phishing* vem crescendo desde 2011, representando 2/3 dos ataques relacionados à espionagem cibernética nos últimos anos. A pesquisa também apresenta que 23% dos destinatários abrem as mensagens de *phishing* e 11% deles acessam os anexos, sendo que 50% das vítimas acessam as páginas falsas na primeira hora de envio [Verizon 2015].

As fraudes e outras formas de ataque envolvendo Engenharia Social estão em constante crescimento, porém existem opções para que as empresas possam testar e avaliar suas equipes. Uma ferramenta livre e acessível para realização dessas análises é o *Social Engineering Toolkit* (SET), um *framework* de *software* que oferece uma gama de ferramentas e opções de testes diferentes, sendo um dos mais populares meios para a realização de testes de Engenharia Social [Watson et al. 1982].

Neste artigo, realizamos uma pesquisa acadêmica sobre Engenharia Social a fim de identificar conceitos, resultados de testes já realizados e análise de ferramentas de apoio à realização de testes. Nesse cenário foi trabalhada uma metodologia de testes baseada em envio de mensagens maliciosas para treinamento e conscientização de usuários, realizando campanhas a fim de identificar pontos fortes e fracos da mesma, visando um método definitivo acessível e de fácil reprodução.

2. Trabalhos Relacionados

O trabalho elaborado por Smith e Toppel (2010) apresenta uma metodologia para elaboração de um teste de Engenharia Social com as diretrizes a serem consideradas, e apresenta os resultados de duas campanhas realizadas em uma companhia, uma com 16 mil gerentes e outra com três mil colaboradores com alto nível de acesso aos servidores da companhia. O estudo apresentou resultados interessantes, como melhor funcionamento do teste quando usados formulários web e o aumento do número de mensagens suspeitas reportadas aos times de resposta à incidentes após os testes, embora incluindo diversas mensagens válidas, indicando insegurança dos usuários em responder, abrir anexos ou clicar em links.

Um estudo de Downs et. al. (2007) realizou um teste em forma de formulário para um grupo de 232 voluntários de uma universidade nos EUA, o qual apresentava tópicos com diferentes tipos de e-mails, *websites*, URLs ou similares, e perguntava aos usuários se confiariam no conteúdo ou se reconheciam indicações de segurança, como o cadeado no navegador. Os resultados do trabalho apresentam duas estratégias no combate ao *phishing*: a necessidade de utilizar métodos para capacitar os usuários na identificação de fraudes, e tornar esses métodos mais compreensíveis aos usuários menos experientes, além de ferramentas de identificação que façam mais do que simplesmente informar o usuário e pedir um “Ok” ou “Cancelar”.

Outro estudo feito na mesma universidade por Kumaraguru et al. (2007) realizou testes com 73 usuários com o objetivo de medir os resultados de treinamentos de conscientização antifraude. Para tal, os participantes foram divididos em três grupos, todos recebendo diversos e-mails falsos em duas etapas. O primeiro grupo recebia e-mails com campanhas antifraude, o segundo grupo recebia treinamento embutido ao *phishing*, e no exato momento em que o usuário caía na fraude, um material educativo era apresentado, e o terceiro grupo não recebia nenhum tipo de treinamento. Os voluntários acreditavam estar participando de um programa sobre testes de

gerenciamento de e-mail, trabalhando durante sete dias em um laboratório que simulava uma mesa de trabalho de um executivo. Eles eram então informados sobre bancos e serviços nos quais o executivo teria cadastro, e lhes era solicitado que tratassem os e-mails recebidos, como em uma situação da vida real. A principal contribuição da pesquisa foi a indicação de que os voluntários que receberam o treinamento embutido ao teste tiveram melhores resultados na identificação de fraudes do que o grupo que recebeu os treinamentos por e-mail. Também foi verificado que o grupo que recebeu treinamento embutido ao teste teve mais atenção ao treinamento e falou mais sobre ele do que os outros grupos.

Um dos maiores estudos de engenharia social já realizados foi o de Mohebzada et al. (2012), o qual realizou dois testes envolvendo mais de 10.000 pessoas, entre funcionários, alunos e ex-alunos de uma universidade nos EUA. O primeiro teste fraudava um e-mail do setor de TI da universidade solicitando a uma troca de senha, o qual em dez dias vitimou um total de 954 usuários, mesmo com o setor de TI enviando um e-mail de alerta. O segundo teste envolvia um falso grupo de pesquisas que pedia voluntários para preencher um formulário com informações pessoais e bancárias, vitimando 220 usuários em 18 horas. O estudo apresentou a falta de atenção dos usuários em relação às mensagens de alerta de fraude, além de desfazer as constantes hipóteses de associar dados como idade e gênero com a possibilidade de ser vítima de fraude.

3. Método Analisado

O método analisado se baseia em um processo de cinco fases: autorização, planejamento, montagem da infraestrutura *web*, montagem e envio dos e-mails e coleta e análise dos resultados. Todo teste de Engenharia Social deve iniciar pela autorização explícita de um gestor, sendo uma boa prática que o cargo do mesmo seja superior ao do alvo com maior nível de hierarquia. A etapa de planejamento então envolve a análise dos alvos e a escolha do serviço/sistema a ser clonado, diretrizes iniciais que irão nortear o restante do teste. A escolha do serviço ou sistema a ser clonado deve dar preferência a páginas que solicitem credenciais de acesso, pois assim é possível diferenciar usuários que apenas acessaram a página dos que realmente foram vítimas da fraude.

A etapa de montagem da infraestrutura *web* envolve uso de um servidor para hospedagem do *phishing*. Nos testes executados foi criada uma máquina virtual Linux no serviço Amazon AWS dentro do plano *Free Tier*, sem custos dentro de determinado nível de uso, sendo que foi possível realizar duas campanhas sem ultrapassar o limite, portanto, sem cobrança. Nesse servidor foi então instalado o *Social Engineering Toolkit* (SET) para clonagem dos *websites* válidos. Para maior credibilidade, no primeiro teste foi utilizado um domínio BR devidamente registrado, com grafia semelhante ao original, ao custo de 30 reais. Para o segundo teste utilizou-se um registro gratuito, tendo, porém, uma URL mais chamativa para identificação do *phishing*. A escolha por um domínio registrado ou gratuito pode ser equilibrada com o nível de dificuldade para identificação da fraude nos demais aspectos do teste. Não utilizar um domínio pode causar problemas com filtros anti-SPAM no envio dos e-mails. Os domínios devem ser então configurados no servidor *web*.

Através do SET é feita então a clonagem dos *websites* verdadeiros, sendo importante no momento da clonagem que o endereço de retorno ser apontado para o domínio criado anteriormente. Para garantia de que não ocorram cruzamentos dos dados que pudessem comprometer a identificação dos resultados da pesquisa, utiliza-se arquivos individuais para cada alvo do teste. A Figura 1 apresenta a estrutura de arquivos utilizada.

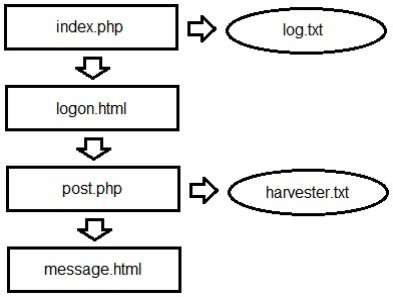

Figura 1. Estrutura de arquivos do *phishing*

O primeiro arquivo, index.php, captura o IP do usuário, a data e a hora do acesso, o endereço de origem (*HTTP Referer*) e as informações do navegador utilizado. As informações são então salvas no arquivo log.txt e o usuário é automaticamente redirecionado para o arquivo logon.html, que é a página clonada com o auxílio do SET. O arquivo logon.html é o *phishing* em si, a página que deve capturar as credenciais do usuário. O meio de validar que um usuário realmente foi capturado seria a captura das credenciais, e o acesso ao sistema/serviço verdadeiro com as mesmas. Esse tipo de teste, porém, entra em conflito com boas práticas de segurança e o fato do *phishing* não utilizar criptografia faz com que as credenciais do usuário vitimado sejam trafegadas em texto claro, sendo, portanto, uma boa prática que própria página do *phishing* apague a senha antes mesmo do envio, capturando apenas o nome de usuário.

Uma vez submetidas as credenciais, o próximo arquivo chamado é o post.php, responsável pela captura das credenciais e das mesmas informações que o arquivo index.php, de forma que seja possível comparar as mesmas e garantir que é o mesmo usuário que acessou e forneceu as informações. As informações do usuário são então enviadas para o arquivo harvester.txt, e o usuário é automaticamente redirecionado para o arquivo message.html. Esse arquivo é uma mensagem para o usuário que cair no teste. Ele deve iniciar com a identificação da empresa e frisar que o usuário foi vítima de um teste devidamente autorizado pela direção da companhia. Em seguida, deve ser solicitado ao usuário que não comente com seus colegas a respeito do teste, evitando assim que seja gerado um alerta que comprometa os resultados. O texto deve então reiterar que a senha do usuário não foi comprometida, e apresentar os objetivos do teste, que são avaliar o nível de capacidade dos colaboradores para identificação de fraudes e, baseado nisso, montar futuras estratégias de treinamento e capacitação. Alertar o usuário que o teste não visa punir quem foi vítima da fraude é uma prática recomendável para evitar situações desagradáveis. Em seguida a mensagem deve mostrar ao usuário as falhas da fraude enviada, mostrando todas as "iscas" deixadas, como o e-mail de origem de um domínio inválido, a URL com grafia incorreta, a ausência de criptografia na página, etc. Esse é o mais importante item do teste, pois é a ferramenta que vai auxiliar

os usuários a fixar os conhecimentos necessários para identificar um ataque verdadeiro. Logo abaixo a mensagem deve oferecer links e outras informações que auxiliem os usuários a descobrir mais sobre o assunto, aprofundando o conhecimento gerado, como políticas ou diretrizes de segurança da empresa. Outra boa prática é a recomendação da Cartilha de Segurança para a Internet do Cert.BR, disponibilizada gratuitamente e numa linguagem acessível para usuários leigos. O final da mensagem deve apontar ao usuário formas de submeter mensagens suspeitas para os responsáveis técnicos que possam avaliar potenciais fraudes, sendo que a notificação dessas tentativas de ataque deve ser uma orientação a todos os usuários.

A fim de criar os arquivos individuais, sugere-se a utilização de numeração nos nomes dos arquivos, por exemplo index1.php, logon1.html, e por diante. É importante lembrar que além de criar os arquivos é necessário alterar o conteúdo dos mesmos, pois possuem links para o próximo arquivo. A codificação errada no conteúdo irá misturar os dados capturados e afetar os resultados do teste. É importante também conferir as permissões de acesso dos arquivos log.txt e harvester.txt para garantir que seja possível gravar as informações capturadas neles.

A próxima etapa é a criação do e-mail. Tendo um domínio, é possível criar um e-mail no mesmo gratuitamente por 30 dias através do serviço Google Apps, o que foi utilizado no primeiro teste realizado. Para o segundo teste foi criado um e-mail simples no serviço Gmail, o qual não tem custos. Para a criação do corpo do e-mail, é ideal seguir um modelo de mensagem original, lembrando-se de utilizar a URL correta para cada usuário. Uma boa prática é realizar alguns testes de envio do e-mail antes de encaminhar aos alvos, a fim de evitar que o mesmo seja identificado por filtros anti-SPAM e bloqueado.

Uma vez enviados os e-mails, os resultados capturados podem ser verificados conferindo o conteúdo dos arquivos log.txt e harvester.txt. Deve-se aguardar um período suficiente para que todos os usuários possam ter visto a mensagem, sendo sugerido cerca de 48 horas. Os usuários que forem vitimados recebem o material informativo no mesmo momento que caírem na fraude, porém quando o teste se der por encerrado, deve ser enviado um e-mail de um endereço válido da empresa para todos os colaboradores no escopo do teste, explicando a realização do teste e apresentando as informações presentes na mensagem vista pelos usuários vitimados. É importante mais uma vez frisar que o objetivo do teste não deve ser o teste em si, mas sim o treinamento e conscientização dos usuários para que sua capacidade de reconhecer fraudes aumente.

4. Testes Realizados

O primeiro teste foi realizado em novembro de 2015, envolvendo um total de 179 usuários com variados níveis de conhecimento divididos em três ondas. O ambiente envolvia colaboradores do setor de TI de uma instituição de nível superior, sendo o teste devidamente autorizado junto aos responsáveis e gestores do setor. A proposta de *phishing* envolveu um clone da página de acesso do serviço *Outlook Web Access* (OWA), utilizada pelos colaboradores como serviço de e-mail, sendo então enviada aos mesmos uma mensagem de correio eletrônico simulando as mensagens reais do serviço enviadas quando a cota de armazenamento da conta está próxima do limite. A Figura 2 mostra o modelo de e-mail enviado aos usuários, sendo que o link *Log in here*

direcionava o usuário para uma página individual. As informações que permitam identificar a instituição foram suprimidas por questões de privacidade.

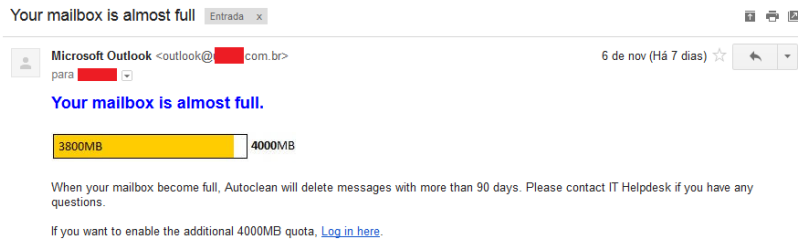

Figura 2. Modelo de e-mail utilizado

A primeira onda foi enviada para um grupo pré-selecionado de cinco usuários de nível técnico mais avançado para se ter uma ideia inicial e obter sugestões sobre a qualidade do teste. A segunda onda por sua vez ocorreu para grupo de 13 usuários, envolvendo diretores, responsáveis de área, líderes de equipe e outros cargos de gestão, servindo como validação para que a direção da instituição autorizasse a última onda de testes. A terceira e última onda então ocorreu com os demais 161 usuários do setor, sendo a grande maioria destes usuários regulares com nível de conhecimento básico a respeito de *phishings*.

O segundo teste foi realizado também em novembro de 2015, em uma única onda com 47 usuários, sendo a grande maioria pessoal técnico. O ambiente envolvia colaboradores dos setores de TI, comercial e administrativo de uma empresa privada da área de tecnologia, sendo o teste devidamente autorizado junto aos responsáveis e gestores do setor. A proposta de *phishing* envolveu um clone da página de serviço de calendário da companhia, utilizado para marcação de reuniões e compromissos, hospedado na plataforma WebCalendar, sendo então enviada aos colaboradores uma mensagem de correio eletrônico simulando as mensagens reais de agendamento de reunião enviadas pelo serviço. A Figura 3 apresenta o modelo de e-mail enviado aos colaboradores. Como o modelo original apresentava uma URL completa, quando se configurava a mesma para direcionar para outra URL, porém mostrando a original, os filtros anti-SPAM identificavam a mensagem como fraude. O recurso utilizado foi substituir a URL no corpo do e-mail por uma imagem da mesma, com o *link* para o *phishing* associado à imagem, sendo assim possível burlar os filtros.

Figura 3. Modelo de e-mail utilizado

5. Análise dos Resultados

O primeiro teste realizado envolveu 179 usuários, sendo destes cinco usuários técnicos e capacitados, 13 usuários gerenciais e 161 usuários com treinamento básico. Numa análise de gênero, 42 eram mulheres e 137 eram homens. Do total do grupo, apenas 12 acessaram a página do *phishing*, sendo uma mulher e 11 homens; quatro usuários técnicos, três gerenciais e cinco usuários padrão. Sete usuários entraram com dados apenas de teste e nenhum foi vitimado, sendo uma possibilidade para tal a opção acordada com os responsáveis para criação das mensagens falsas em inglês, enquanto o sistema de e-mail utilizado era em português. Não há como garantir se o real motivo da ausência de vítimas foi esse, porém a utilização de fraudes excessivamente fáceis de serem percebidas irão prejudicar os resultados esperados no trabalho. Também foi verificado que 13 usuários reportaram a tentativa de fraude ao setor responsável. Depois do evento, através de uma amostra de entrevistas individuais, houve a confirmação de que o teste de fato ampliou o nível de preocupação dos usuários, que reportaram estar mais atentos a partir daquele momento.

O segundo teste envolveu 47 usuários, sendo destes 42 usuários técnicos e capacitados, e cinco usuários padrão, porém com treinamentos na área de segurança. Numa análise de gênero, sete eram mulheres e 40 eram homens. Do total do grupo, 14 acessaram a página do *phishing*, sendo todos do sexo masculino e com nível técnico de conhecimento. Desses, dez foram vitimados pelo teste, incluindo dois líderes de equipe técnica. Nenhum usuário utilizou os canais formais da empresa para notificar o incidente, mesmo os que perceberam a fraude. A equipe de segurança da informação elogiou a montagem e análise dos resultados do teste, sendo que os testes passarão a ser realizados periodicamente como parte das campanhas de segurança da empresa da companhia, e o teste e treinamentos específicos serão adicionados aos treinamentos periódicos já realizados com as equipes.

Uma situação observada nos testes chamou a atenção durante a análise dos resultados. Um usuário que foi vítima da fraude cerca de cinco minutos após ter submetido suas credenciais submeteu em sequência os nomes de usuário verdadeiros de outros sete colaboradores, inclusive de gestores e executivos da empresa, possivelmente temendo os resultados do teste. Essa atitude traz à tona a importância de conscientizar os usuários quanto aos objetivos do teste, que não envolvem a punição, e sim o treinamento. Também foi assim validada a importância da utilização de arquivos individuais para cada alvo do teste, sendo possível diferenciar a identidade e informações fornecidas por cada usuário, com menores riscos de cruzamento de informações ou tentativas de manipular os resultados do teste por parte dos usuários. Agradecemos às organizações que nos autorizaram a execução desse experimento em ambientes reais e representativos, sem os quais esse trabalho não seria possível.

5. Considerações Finais

O presente artigo utilizou uma metodologia para realização de testes de Engenharia Social baseada em envio de e-mail malicioso. Utilizando a estrutura sugerida foi possível realizar os dois testes com orçamentos mínimos e com uma configuração tecnicamente simples. A montagem da estrutura não apresentou dificuldades, e o retorno recebido dos envolvidos foi positivo.

Uma das propostas iniciais foi a modificação do código do SET para automatizar a montagem da estrutura de arquivos. Após análise do código e tentativas sem sucesso foi concluído que a abordagem não atenderia aos objetivos do trabalho. Optou-se então por utilizar apenas o mecanismo de clonagem, que funcionou corretamente nos testes realizados. Outra situação identificada foi um problema na captura das credenciais quando o *phishing* padrão criado pelo SET era utilizado, não salvando as informações.

A configuração da captura das informações dos usuários que tanto apenas acessavam a página quanto dos que submetiam as credenciais em arquivos individuais permitiu a correta identificação dos usuários e o foco de futuros treinamentos para ambas as situações. O teste também mostrou a importância de campanhas que incentivem os usuários a reportarem mensagens suspeitas, visto que um ataque direcionado deve gerar um alerta imediato nas equipes de reposta à incidentes.

Dentre trabalhos futuros, além das oportunidades já citadas, destaca-se (I) o desenvolvimento de uma aplicação que automatize a criação e montagem da infraestrutura necessária para realização do teste; e (II) a realização de mais testes seguindo o modelo apresentado, inclusive nos mesmos locais a fim de verificar melhorias nos níveis de identificação de fraudes por parte dos colaboradores.

Referências

Darwish, Ali, El Zarka, Ahmed e Aloul, Fadi. (2012) “Towards understanding phishing victims' profile”, http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6454454.

Downs, Julie S., Holbrook, Mandy e Cranor, Lorrie Faith. (2007) “Behavioral response to phishing risk”, http://dl.acm.org/citation.cfm?id=1299015.1299019.

Hadnagy, Christopher. (2011) Social Engineering. Wiley Publishing.

Kumaraguru, Ponnurangam, Rhee, Yong, Sheng, Steve, Hasan, Sharique, Acquisti, Alessandro, Cranor, Lorrie Faith e Hong, Jason. (2007) “Getting Users to Pay Attention to Anti-Phishing Education: Evaluation of Retention and Transfer”, http://dl.acm.org/citation.cfm?id=1299015.1299022&coll=DL&dl=ACM.

Mohebzada, Jamshaid G., El Zarka, Ahmed, Bhojani, Arsalan H. e Darwish, Ali. (2012) “Phishing in a university community: Two large scale phishing experiments”, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6207742.

PwC. (2015) “Key Findings from the Global State of Information Security Survey 2016”, http://www.pwc.com/gx/en/issues/cyber-security/information-security-survey/download.html.

Smith, Allen M. e Toppel, Nancy Y. (2010) “Use of Spear Phishing Exercises do Increase Security Awareness”, http://cisse.info/resources/archives/category/14-papers?download=165:1716-2010.

Verizon. (2015) “2015 Data Breach Investigations Report”, http://www.verizonenterprise.com/DBIR/2015/.

Watson, Gavin, Mason, Andrew e Ackroyd, Richard. (2014) Social Engineering Penetration Testing: Executing Social Engineering Pen Tests, Assessments and Defense. Elsevier.