Engenharia social: explorando o efeito halo para obter acesso físico não autorizado em empresas de cartões de pagamento

Guilherme Gattino*, Jéferson Campos Nobre†
Centro de Informática
Segurança da Informação
UNISINOS
São Leopoldo, Rio Grande do Sul
Email: *ggattino@gmail.com, †jcnobre@unisinos.br

***Resumo*—Diversas organizações investem muito dinheiro em controles tecnológicos para proteger as suas informações e controlar os acessos às suas dependências. Porém, conhecer os fatores humanos e saber como induzir falhas nesse aspecto, aprimora as chances de sucesso de um ataque de engenharia social e permite que esses controles sejam facilmente burlados. Este artigo apresenta o uso do viés cognitivo efeito halo como uma maneira de aprimorar ataques de engenharia social que visam trespassar a autorização de visitantes em organizações com os controles de segurança do Padrão de Segurança de Dados do Setor de Cartões de Pagamento implementados.**

I. Introdução

A informação tem uma importante função para as mais diversas corporações, pois é através da informação que elas detêm os seus planos, projetos, valores e demais dados que podem definir o seu sucesso ou o seu fracasso. Proteger a informação contra acesso indevido e evitar que ela seja divulgada e/ou obtida de forma não autorizada, exige a inserção de controles e procedimentos que visam garantir os atributos básicos de segurança da informação, seja através de recursos tecnológicos e/ou processos de negócio.

Esses controles perdem a sua eficiência quando os aspectos comportamentais e psicológicos humanos não fazem parte do escopo da segurança da informação, o que normalmente acontece na maioria das organizações [1]. A negligência desses aspectos pode facilitar que uma pessoa mal intencionada consiga explorar as suas falhas e obter acesso não autorizado às informações ou até mesmo ao interior de uma organização, trespassando as mais poderosas ferramentas tecnológicas de proteção e os seus processos para controle da segurança da informação [2].

A falta de proteção aos aspectos humanos expõe as organizações a riscos reais, pois uma pessoa que possua conhecimento e sucesso em explorar técnicas de engenharia social pode obter e divulgar informações sensíveis e de alto valor para a corporação, além da possibilidade de obter acesso indevido em ambientes físicos e em redes de computadores. Isso só é possível pela existência de inúmeros aspectos psicológicos que podem ser utilizados para induzir pessoas ao erro, forçando-as a divulgar informações, realizar ações que normalmente não fariam e até mesmo mudar as suas opiniões.

II. A engenharia social e o aspecto humano

A metodologia mais comum utilizada pelas organizações para proteger as suas informações é a aplicação de controles tecnológicos, como *hardwares* e *softwares* específicos, investindo altos valores financeiros. Além dessas medidas tecnológicas, algumas organizações decidem por utilizar processos de negócio que visam garantir essa proteção, buscando aplicar políticas, procedimentos, processos e estruturas organizacionais com o objetivo de manter ainda mais efetivos esses controles tecnológicos de segurança da informação.

Mesmo com um alto valor de investimento em controles de proteção da informação, é possível que esses controles sejam simplesmente burlados quando outro aspecto, muitas vezes negligenciado, é explorado por entidades mal intencionadas: o aspecto humano da segurança da informação. Um indivíduo que tenha conhecimento de como forçar falhas dentro desse contexto e explorá-las, pode simplesmente trespassar todos os controles técnicos e obter acesso às informações consideradas confidenciais ou até mesmo obter acesso físico em ambientes considerados críticos [2].

A. Engenharia social

A engenharia social é definida como a ciência de manobrar habilmente as pessoas através do uso de erros e fraquezas humanas (*i.e.*, vieses cognitivos), para que realizem determinadas ações ou divulguem informações confidenciais [4]. O seu objetivo é obter acesso físico ou lógico não autorizado, além de obter informações sensíveis através de técnicas de manipulação de pessoas ou através de técnicas que as induzam a erros em seus julgamentos e tomadas de decisão.

A forma como as pessoas pensam é um aspecto humano crucial para realizar um ataque de engenharia social. O cérebro humano, quando recebe qualquer informação, tem de processá-la internamente para que seja possível tomar decisões, realizar previsões e julgamentos. Esses processos funcionam na maioria das circunstâncias, porém em certas ocasiões levam a erros sistêmicos [5] e padrões de desvio no julgamento que levam a percepções distorcidas, julgamento impreciso e interpretações ilógicas, conhecidos como viés cognitivo [6].

Com a existência de falhas no processamento de informação no cérebro humano é possível forçar esses erros

para aumentar a probabilidade de sucesso de um ataque de engenharia social. Assim, um ou mais vieses cognitivos podem ser explorados para que um acesso físico seja permitido, mesmo esse não sendo autorizado. Em muitas organizações o controle de acesso físico é realizado através de recepcionistas e guardas patrimoniais, porém isso não impede que outros controles possam ser utilizados para tentar impedir acessos físicos. Em empresas que trabalham com dados de cartão, por exemplo, para controle de acesso físico devem ser aplicadas regras específicas para visitantes (*e.g.*, autorização prévia do visitantes antes de sua entrada na empresa, com o objetivo de reduzir as chances de que um acesso não autorizado seja efetuado) [7].

B. Viés cognitivo e a segurança da informação

Muitas vezes as pessoas tomam decisões baseadas na crença de probabilidades de que um determinado evento com conclusão incerta possa ocorrer [8]. Como as pessoas possuem um conjunto limitado de conhecimentos que podem ser aplicados para a resolução de problemas [9], muitas tarefas complexas para calcular probabilidades e previsões de valores são reduzidas para simples julgamentos. Esses julgamentos são baseados em dados que possuem uma eficácia limitada, ao qual são processados de acordo com princípios heurísticos. A confiança nesses julgamentos leva a erros sistemáticos, que ocasionam desvios dos mesmos, em interpretações ilógicas, distorções perceptivas ou ações irracionais [8]. Esses erros são conhecidos como viés cognitivo.

Dentro do contexto da segurança da informação, o viés cognitivo pode ser explorado em alvos humanos, induzindo-os a abrir brechas de segurança, como a divulgação de informações confidenciais, liberação de acessos indevidamente ou a inserção de discos removíveis em uma unidade USB. Um engenheiro social com o conhecimento de como induzir pessoas a desvios nos padrões de julgamentos, percepções e tomadas de decisão (*i.e.*, erro), e com a capacidade de explorar um ou mais vieses cognitivos, pode estar facilitando todo um processo para obter informação ou acesso físico em um ambiente.

Um viés cognitivo que possibilita induzir pessoas ao erro é o efeito halo, que em uma definição geral é a influência que uma avaliação global possui em uma avaliação dos atributos individuais de uma pessoa [10]. Como o efeito halo distorce a avaliação de atributos distintos com base em uma impressão geral, uma pessoa que tenha maiores cuidados com a sua aparência física, ou seja, que sua aparência física seja agradável para o alvo, estes atributos serão superavaliados pelo expectador. O mesmo pode ocorrer com características inversas, onde uma aparência desagradável ao alvo pode fazer com que outros atributos sejam subavaliados.

A aparência física de uma pessoa é a característica mais óbvia e acessível para outras pessoas em uma interação social, ao qual permite que, psicologicamente, um indivíduo crie determinadas expectativas simplesmente por conhecer a aparência de outro. Isto ocorre porque existe uma correlação entre características pessoais e a aparência [11]. Além dessa correlação com características pessoais, os estereótipos culturais sobre os tipos de personalidades apropriadas para caracterizar beleza ou feiura podem moldar as características desses indivíduos, como em culturas onde determinados atributos (*e.g.*, sinceridade, nobreza e honestidade) são definidos como aceitação ou atratividade social.

Quando há a necessidade de tomar decisões e julgamentos, as pessoas utilizam um conjunto de crenças e atitudes que as influenciam, o que eleva a probabilidade de ocorrerem distorções quando essas pessoas estão presenciando algum evento inconsistente ao seu pensamento [12]. A aparência física de uma pessoa é a característica mais óbvia e acessível para outras pessoas em uma interação social, ao qual permite que, psicologicamente, um indivíduo crie determinadas expectativas simplesmente por conhecer a aparência do outro.

Portanto, quando o efeito halo ocorre, quando um indivíduo enxerga e determina, seja através de padrões pessoais ou culturais, que outro indivíduo é agradável fisicamente, é bastante provável que esta característica seja extrapolada e estendida para outras características pessoais e independentes [13]. Dessa forma, o efeito halo pode ser explorado para obter acesso físico não autorizado, forçando que uma determinada característica física (*i.e.*, aparência física) seja estendida para outras características independentes (*e.g.*, hierarquia ou confiabilidade) e um possível bloqueio de acesso por autorização seja trespassado.

III. Experimento

Para analisar e comprovar que o efeito halo pode ser utilizado para burlar especificamente o controle *9.3.1* do PCI-DSS, que estabelece um processo para autenticação e autorização de visitantes, foi realizado um experimento. Esse experimento tem uma abordagem mais próxima aos experimentos da área da psicologia e busca analisar, através da tentativa de extrapolar uma avaliação da aparência física para outras características, se o efeito halo pode ou não influenciar na decisão de pessoas para permitir um acesso sem autorização. Como é bastante difícil simular em um ambiente controlado todos os fatores reais de uma organização certificada pelo PCI-DSS e que possua um controle efetivo de autorização de visitas, para a realização do experimento foi utilizado um questionário para obtenção de dados.

A. Metodologia

O experimento iniciou-se com a seleção de duas pessoas, com boas expressões faciais, para serem utilizadas como modelos onde duas fotografias de cada uma dessas pessoas foram capturadas. Para a primeira fotografia, os modelos foram instruídos a criar uma expressão de autoridade e arrogância, além de serem vestidos com roupas mais sociais, mais próximas de que um alto executivo utilizaria. Para a segunda fotografia, os modelos foram instruídos a se mostrarem com expressões faciais mais carismáticas e utilizaram roupas mais informais.

Para cada uma das fotografias, um conjunto de questionamentos foi formulado com o objetivo de classificar os atributos da pessoa nessa fotografia. Foram classificadas as características da aparência física, a confiabilidade, o conhecimento, a simpatia e a influência hierárquica, baseando-se apenas na fotografia apresentada. Após esta avaliação foi solicitado ao voluntário do experimento se ele permitiria ou não que a pessoa obtivesse acesso ao interior da organização. Em seguida, o efeito halo foi medido, ao classificar o grau

de influência que cada característica teve para a tomada de decisão. Antes dos questionamentos, o voluntário foi instruído sobre o controle *9.3.1* do PCI-DSS, através da leitura de um texto previamente preparado, e explicado que o acesso de visitantes só é permitido após uma autorização formal de algum funcionário da empresa.

Para responder os questionamentos, foram selecionados 30 voluntários, os quais foram divididos em dois grupos com base em suas profissões. O primeiro grupo, definido como grupo A, foi composto por 15 voluntários especializados em segurança da informação, o que inclui profissionais da área e estudantes do curso de Segurança da Informação da Universidade do Rio do Vale dos Sinos (UNISINOS). O segundo grupo, definido como grupo B, foi composto por profissionais de segurança patrimonial, o que inclui porteiros, guardas patrimoniais e recepcionistas.

B. Resultados

Os resultados apresentados através da coleta de dados deste experimento são mostrados através de gráficos, com o objetivo de um melhor entendimento de seus significados. Foram realizadas várias séries de comparações buscando identificar a ação do efeito halo em cada um dos modelos fotografados.

A Figura 1 apresenta um gráfico com o total das classificações realizadas pelos voluntários ao analisar as fotografias com expressão autoritária. Este gráfico é a soma de todos os resultados obtidos com as fotografias com a expressão autoritária com os dois modelos. Na análise dos dados dessa soma, percebeu-se que há uma grande quantidade de voluntários que avaliou o autoritarismo com um grau mediano de características. Destacaram-se as características de confiabilidade, simpatia e influência hierárquica que foram avaliadas com esses valores por 60% dos voluntários.

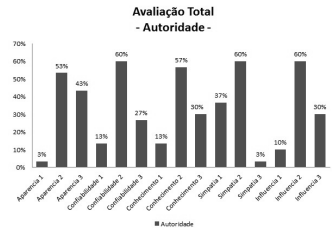

Figura 1. Avaliação total das características dos modelos com expressão autoritária

O próximo gráfico, ilustrado na Figura 2, mostra a soma das classificações realizadas pelos voluntários ao analisar as fotografias com a expressão carismática. Com essa análise, foi possível identificar que a grande maioria das características foram classificadas como medianas (valor equivalente a 2), com exceção da simpatia que foi classificada por 63% dos voluntários com um valor alto.

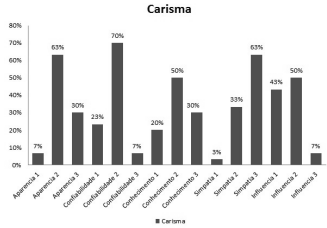

Figura 2. Avaliação total das características dos modelos com expressão carismática

Foi realizada análise para verificar a quantidade de voluntários que permitiu que os modelos nas fotografias acessassem a organização sem uma autorização formal. Mesmo sendo instruídos a solicitar uma autorização prévia antes de permitir o acesso, 50% de todos os voluntários permitiu que o modelo autoritário entrasse na organização. Da mesma maneira, 37% dos voluntários permitiram o acesso do modelo carismático sem uma autorização. É importante notar que, mesmo que em nenhuma das duas expressões as avaliações de suas características tenham sido avaliadas com um valor alto, houve voluntários que permitiram o acesso. A Figura 3 identifica em forma de gráfico a porcentagem de acessos permitidos sem autorização em comparação por tipo de expressão (autoridade *versus* carisma).

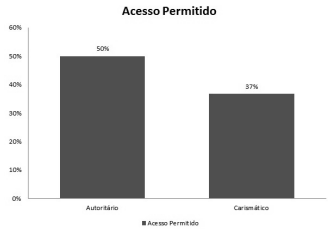

Figura 3. Avaliação total dos acessos permitidos, comparados por expressão (autoritário *versus* carismático)

Para analisar a eficácia de pessoas especializadas em segurança da informação e que possuam determinado conhecimento dos riscos em que a empresa pode estar exposta quando não há o cumprimento das normas impostas, foi realizada uma comparação entre os dois grupos de voluntários. A comparação foi realizada levando em consideração o total de acessos permitidos ao somarem-se os resultados de acessos permitidos das duas expressões dos modelos. Como o esperado, o gráfico da Figura 4 mostra que 57% dos voluntários especializados em proteção patrimonial (grupo B) permitiram o acesso indevido ao interior da empresa, mesmo que uma instrução fosse ordenada antes do questionário ser iniciado. Mas é surpreendente que 30% dos voluntários especializados em segurança da informação (grupo A) permitiram que o acesso fosse realizado.

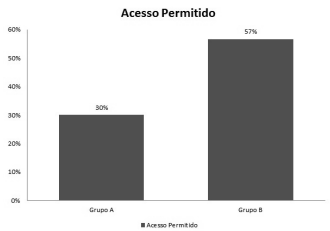

Figura 4. Total de acessos permitidos, comparados por grupo

Finalmente, foi realizada análise da influência que as características avaliadas pelos voluntários tiveram para determinar a sua decisão em permitir ou de não permitir o acesso físico. Para isso, foram consideradas as respostas marcadas com o valor 2 e 3 para cada uma das características, visto que apenas o valor 1 representa nenhuma influência. O gráfico representado na Figura 5 indica que há uma significativa influência da autoridade na tomada de decisão dos voluntários para permitir um acesso. As fotografias de modelos com expressão autoritária passaram confiança para 77% dos voluntários, assim como 67% tiveram suas decisões influenciadas pela aparência e influência julgadas apenas por essas fotografias.

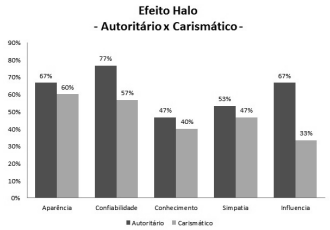

Figura 5. Efeito halo medido por característica avaliada

De acordo com os gráficos apresentados e com os resultados obtidos neste experimento, ficou evidenciado que o efeito halo pode ser explorado para obter acesso físico não autorizado, mesmo que seja exigido um controle de aprovação para liberação deste acesso. Dessa forma, uma pessoa mal intencionada que consiga fazer com que sua aparência seja mais autoritária e que consiga refletir um padrão de vestimentas utilizado pela alta direção da empresa, terá maiores probabilidades de trespassar por esse controle.

IV. Conclusão

Apesar da existência de padrões de segurança das informações e desses padrões apresentarem uma ampla gama de controles, que nem sempre são técnicos, existem chances consideráveis para que pessoas cometam erros comuns na tomada de suas decisões, em suas memórias, em suas crenças e nos seus comportamentos. Estes erros, conhecidos como viés cognitivo, ocorrem em determinadas situações e quando confrontados com o objetivo de obter informações ou acesso não autorizado trespassam os controles técnicos e os mais tecnológicos sistemas de proteção de acesso. Com isso, controles muito bem implementados podem simplesmente não ter o efeito defensivo esperado, pois a falha explorada foi humana, foi um viés cognitivo.

O experimento realizado obteve êxito em demonstrar que o viés cognitivo efeito halo pode ser explorado para burlar um processo de autorização de visitantes e ficou evidente que pessoas que não possuem uma consciência ou especialidade com a segurança da informação são mais suscetíveis a serem enganadas pelas consequências do efeito halo. A aparência física e as expressões faciais são as principais características que levam à exploração desse viés cognitivo, fazendo com que o alvo estenda o seu julgamento nessas características para outras características distintas. Explorando o efeito halo em conjunto com um pretexto para a entrada na organização (*e.g.*, fazendo-se passar por um técnico de informática ou um consultor) as chances de sucesso são ainda mais aprimoradas.

Referências

[1] Y. Lafrance. (2004, Fev.) *Psychology: a precious security tool*, [Online]. Disponível em http://www.sans.org/reading-room/whitepapers/engineering/psychology-precious-security-tool-1409. [Acesso em 18 de Março de 2012]

[2] I. Mann. (2011) *Engenharia social*. Série Prevenção a Fraudes. São Paulo: Blucher.

[3] T. Thornburgh. (2004, Set.) *Social engineering: the dark art*. In *Of the first Annual Conference on Information Security Curriculum Development*. Kenessaw, Georgia. Nova Iorque: ACM. pp.133-135.

[4] A. Thapar. (2007, Jun) *Social engineering: an attack vector most intricate to tackle*. [Online]. Disponível em http://www.infosecwriters.com/text_resources/pdf/Social_Engineering_AThapar.pdf. [Acesso em 16 de Março de 2012]

[5] G. Gigerenzer. (1991) "How to make cognitive illusions disappear: beyound heuristics and biases", *European Review of Social Psychology*. Vol. 2. pp.83-115.

[6] J. Taylor. (2011, Jul). *Understanding cognitive bias*. [Online]. Disponível em https://www.huffingtonpost.com/dr-jim-taylor/cognitive-biases_b_896421.html. [Acesso em Agosto de 2012]

[7] PCI Security Standards Council. (2010, Out.) *Navegando pelo PCI DSS: conhecer a intenção dos requisitos — Versão 2.0*.

[8] A. Tversky e D. Kahneman. (1974) "Judgement under uncertainty: heuristics and biases", *Sciences*. Vol. 185. No. 4157. pp.1124-1131.

[9] J. Firmino e T. Broto.(2009) "Raciocínio, heurísticas e resolução de problemas: um diálogo teórico-conceitual", *Mosaico: estudos em psicologia*. Vol. 3. No. 1. pp.1-12.

[10] R. Nisbett e T. Wilson. (1977) "The halo effect: evidence for unconscious alteration of judgments", *Journal of Personality and Social Psychology*. Vol. 35 No. 4. pp.250-256.

[11] K. Dion, E. Berscheid e E. Walster. (1972) "What is beautiful is good". *Journal of Personality and Social Psychology*. Vol. 24. No. 3. pp.285-290.

[12] L. Leuthesser, C. Kohli e K. Harich. (1995) "Brand equity: the halo effect measure", *European Journal of Marketing*. Vol. 29. pp.57-66.

[13] C. Hadnagy. (2011). *Social engineering: the art of human hacking*. Indianapolis: Wiley Publishing.