Phishing

André Baião, Gonçalo Barradas, Guilherme Grilo
48092, 48402, 48921

março de 2022

Resumo

Atualmente a utilização da internet é generalizada e essencial ao nosso dia-a-dia, havendo por isso cada vez mais partilha de informação online por parte do utilizador.

Como resultado desta enorme partilha de informações e transações financeiras existe uma maior vulnerabilidade ao cibercrime. O phishing é uma das formas mais eficazes e utilizadas de crimes cibernéticos, sendo utilizada contra utilizadores individuais, empresas e agências corporativas ou governamentais [Alkhalil et al., 2021].

Nos últimos tempos temos assistido a cada vez mais fraudes online, ataques a sistemas de grandes empresas, bem como acesso a informações confidenciais de empresas, bancos etc,.. existem casos bastante mediáticos tanto no nosso país como noutros que incluem este tipo de crimes, sendo por isso essencial arranjar soluções para proteger as empresas e as pessoas deste tipo de crimes. O objetivo deste trabalho é avaliar as metodologias que estão implementadas e as que estão a ser desenvolvidas para combater o phishing.

O processo normalmente utilizado num ataque phishing consiste no envio de conteúdo de engenharia social ou links fraudulentos através de e-mail, mensagens instantâneas ou redes sociais. O usuário abre os links que dão acesso a sites phishing, que visualmente se assemelham a sites conhecidos de uma determinada marca. O usuário insere informações confidenciais, que são posteriormente utilizadas para acesso não autorizado a contas bancárias, e-mail, possibilitando também o roubo de identidade [Karamagi, 2022][Liu et al., 2021] [Minocha and Singh, 2022].

Solucionar o phishing é um grande desafio, sendo necessário haver avanços na área da segurança, software e eletrónica, sendo também essencial haver investimento em cybersegurança por parte do governo e das empresas[Lacey et al., 2015]

Já existem várias listas disponíveis com sites que foram relatados como phishing. Para identificar sites phishing, têm sido desenvolvidos vários estudos, essencialmente na área da Inteligência Artificial, através da utilização de CNN, LSTM [Alshehri et al., 2022][Xiao et al., 2021].

A identificação de sites phishing, tem por base a analise de 2 tipos de recursos: os recursos relacionados à semelhança da página web, e os recursos que implementam a função de roubo. Dentro da semelhança das páginas web normalmente é analisado o logótipo da empresa e os respetivos direitos de autor, links recebidos e informações do sistema de nome de domínio (DNS). Os recursos de roubo referem-se a recursos associados ao roubo de informações confidenciais, como números de contas e senhas, incluindo "recurso de formulário", "recurso de submissão", "recurso de senha", "recurso https". No "recurso de formulário"por exemplo, é avaliado se a tag form se encontra dentro da página web ou não, através da análise dos valores 0 ou 1 inscritos no código [Liu et al., 2021].

Referências

[Alkhalil et al., 2021] Alkhalil, Z., Hewage, C., Nawaf, L., and Khan, I. (2021). Phishing attacks: A recent comprehensive study and a new anatomy. *Frontiers in Computer Science*, 3. https://doi.org/10.3389/fcomp.2021.563060.

[Alshehri et al., 2022] Alshehri, M., Abugabah, A., Algarni, A., and Almotairi, S. (2022). Character-level word encoding deep learning model for combating cyber threats in phishing url detection. *Computers Electrical Engineering*, 100:107868. https://doi.org/10.1016/j.compeleceng.2022.107868.

[Karamagi, 2022] Karamagi, R. (2022). A review of factors affecting the effectiveness of phishing. *Computer and Information Science*, 15(1):20 – 31. https://doi.org/10.5539/cis.v15n1p20.

[Lacey et al., 2015] Lacey, D., Salmon, P., and Glancy, P. (2015). Taking the bait: A systems analysis of phishing attacks. *Procedia Manufacturing*, 3:1109–1116. https://doi.org/10.1016/j.promfg.2015.07.185.

[Liu et al., 2021] Liu, D.-J., Geng, G.-G., Jin, X.-B., and Wang, W. (2021). An efficient multistage phishing website detection model based on the case feature framework: Aiming at the real web environment. *Computers Security*, 110:102421. https://doi.org/10.1016/j.cose.2021.102421.

[Minocha and Singh, 2022] Minocha, S. and Singh, B. (2022). A novel phishing detection system using binary modified equilibrium optimizer for feature selection. *Computers Electrical Engineering*, 98:107689. https://doi.org/10.1016/j.compeleceng.2022.107689.

[Xiao et al., 2021] Xiao, X., Xiao, W., Zhang, D., Zhang, B., Hu, G., Li, Q., and Xia, S. (2021). Phishing websites detection via cnn and multi-head self-attention on imbalanced datasets. *Computers Security*, 108:102372. https://doi.org/10.1016/j.cose.2021.102372.