Uma extensão para o VSCode que utiliza o ChatGPT como ferramenta de apoio ao desenvolvimento de software seguro

Mustafa Ribeiro de Almeida Neto
Bacharelado em Sistemas de Informação
Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (Cefet-RJ)
Nova Friburgo, Brasil
mustafa.neto@aluno.cefet-rj.br

Nilson Mori Lazarin
Bacharelado em Sistemas de Informação
Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (Cefet-RJ)
Nova Friburgo, Brasil
nilson.lazarin@cefet-rj.br

ABSTRACT

This paper proposes the adoption of Generative Artificial Intelligence (GAI) in order to reduce vulnerabilities on the software development process. To analyze this approach, we present a VSCode's extension for the ChatGPT. Furthermore, we proceed a qualitative evaluation of the GAI's suggests to fix various vulnerable codes. The observed results suggest that this approach can be promising, bringing certainty to developers when dealing with software security issues.

KEYWORDS

Cybersecurity, IDE, Generative Artificial Intelligence

1 INTRODUÇÃO

A informação é um ativo importante, têm valor para as organizações e, consequentemente, requerem proteção contra vários riscos. A Segurança de Sistemas da Informação (SSI), por sua vez, é obtida pela implementação de um conjunto adequado de controles, incluindo políticas, processos, procedimentos, estrutura organizacional e funções de software e hardware [1].

De acordo com um relatório da Trend Micro, o Brasil é o segundo país mais vulnerável a ataques de hackers e teve mais de 7 bilhões de ameaças no primeiro semestre de 2023 [2]. Muitos desses ataques buscam explorar falhas nos softwares, por isso, uma parte significativa do ciclo de vida de um sistema é a revisão do código-fonte em busca de vulnerabilidades [3]. Entretanto, ao considerarmos o fator humano no processo de desenvolvimento e manutenção do software, os egressos de cursos superiores consideram ter recebido uma formação inadequada para lidarem com segurança de sistemas da informação [4, 5].

Por outro lado, a Inteligência Artificial (IA) tem transformado o dia-a-dia das pessoas e organizações, revelando-se uma ferramenta valiosa capaz de auxiliar em diversas tarefas [6]. Em se tratando de SSI, a IA pode ser aplicada de diversas formas, seja na análise de tráfego de rede, seja na análise de logs, no reconhecimento de arquivos maliciosos, ou ainda na busca por vulnerabilidades [7, 8].

Neste contexto, este trabalho apresenta uma abordagem de integração do ambiente de desenvolvimento com Inteligência Artificial Generativa, buscando apoiar os desenvolvedores em tempo de design, de forma que estes possam produzir códigos mais seguros. Para analisar a viabilidade da abordagem, foi desenvolvida uma extensão que integra o Visual Studio Code (VSCode) ao ChatGPT e foi realizada uma análise qualitativa, através de um estudo de caso onde foram analisadas as sugestões recebidas, quando foram submetidos códigos-fonte reconhecidamente vulneráveis.

2 METODOLOGIA

Para avaliarmos a possibilidade do uso de IA para análise de código-fonte em tempo de design, foram escolhidos o VSCode e o ChatGPT. O VSCode foi escolhido por ser um dos ambientes de desenvolvimento mais utilizados no mundo [9] e por permitir a criação de extensões que o personalizam e o aprimoram, incluindo configurações, recursos ou usos extras para ferramentas existentes. O ChatGPT (gpt-3.5-turbo) foi escolhido por ser uma ferramenta de processamento de linguagem natural, que utiliza uma rede neural treinada com milhões de textos da internet e pode ser utilizado para várias finalidades, tais como chatbots, geração automática de conteúdo, tradução automática, entre outras [10].

Na Figura 1 é apresentada a abordagem proposta. O desenvolvedor, em tempo de design, seleciona o trecho de código que deseja analisar e executa a extensão. A extensão prepara o prompt e a conexão com a IA. Após o retorno da IA, a extensão exibe as recomendações ao desenvolvedor.

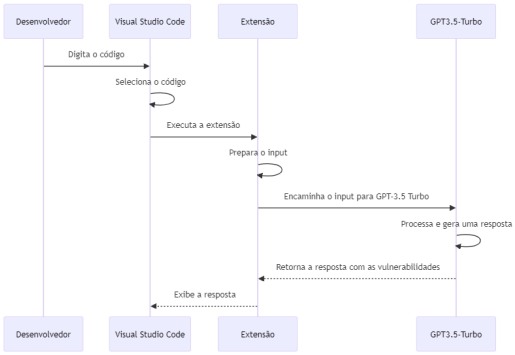

Figure 1: Interação entre Desenvolvedor e a Extensão de Análise de Código.

2.1 Implementação

A extensão opera mediante a seleção de código realizada pelo desenvolvedor no VSCode, utilizando essa seleção para formar o prompt (entrada) que será enviado ao ChatGPT. O prompt é construído com a instrução: "*Detecte as vulnerabilidades no código e explique, se houver alguma*", seguido do trecho de código selecionado.

Internamente, a extensão utiliza uma chave de API do OpenAI para acessar o modelo GPT-3.5-turbo. O código selecionado é encapsulado em uma mensagem, então enviada para a API do OpenAI.

Após receber a resposta da API, a extensão cria uma visualização em formato de painel no VSCode.

A extensão está disponível para download[1] e um vídeo tutorial[2] apresenta seu uso. Após instalada, basta selecionar o trecho de código que deseja analisar e executar o comando "*Buscar vulnerabilidades*" na paleta de comandos. Na Figura 2 é apresentado o uso da extensão para analisar um trecho de código-fonte vulnerável.

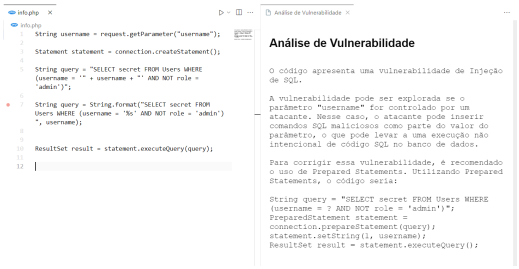

Figure 2: Análise de vulnerabilidade em um trecho de código utilizando a extensão desenvolvida.

3 AVALIAÇÃO QUALITATIVA

Buscando avaliar o desempenho da abordagem proposta, foram realizados dois experimentos utilizando a extensão desenvolvida. No primeiro, submetemos uma coleção[3] de trechos de códigos-fonte reconhecidamente vulneráveis. Ao todo foram testados 34 diferentes tipos de vulnerabilidades. O objetivo deste experimento é avaliar a precisão do uso de IA Generativa na identificação de vulnerabilidades em tempo de design. No segundo, foram analisados quatro trechos de códigos-fontes vulneráveis utilizados por Costa et. al, 2018 ([4]), neste experimento esperamos comparar capacidade de identificação de vulnerabilidades da IA Generativa em relação ao resultado obtido por desenvolvedores pleno e sênior no estudo.

No primeiro experimento, a abordagem apresentou uma eficácia de 88,24% (identificando corretamente 30 das 34 vulnerabilidades apresentadas). Vale ressaltar que, mesmo nos casos em que não foram identificadas vulnerabilidades, em dois deles (*Code Execution* e *PostMessage Security*) a IA Generativa ofereceu considerações pertinentes para aprimorar a segurança do código. Apenas nos casos de *Connection String Injection* e *Sensitive Data Exposure* não foram identificadas vulnerabilidades.

No segundo experimento, a IA Generativa apresentou uma eficácia de 100% na identificação das ameaças contidas no código-fonte. Esse desempenho destaca o valor significativo da abordagem proposta, especialmente considerando que apenas 61% dos desenvolvedores na pesquisa conseguiram identificar vulnerabilidades como *Session Hijacking*.

Nesses experimentos, as recomendações da IA Generativa revelaram-se eficazes na identificação das vulnerabilidades e pertinentes na sugestão de soluções, contribuindo de maneira satisfatória para o entendimento e resolução nos códigos apresentados.

4 DISCUSSÃO

Em um cenário de crescente ameaça cibernética, este trabalho apresentou uma abordagem que busca facilitar a segurança da informação no desenvolvimento de software, através da integração do ambiente de desenvolvimento com a IA Generativa. Foi desenvolvida uma extensão de integração do VSCode com o ChatGPT e mediante uma avaliação qualitativa revelou-se promissora a identificação e análise de vulnerabilidades em diversos códigos.

Os resultados obtidos ao analisar o repositório *Vulnerable Code Snippets* destacaram a eficácia da abordagem, não apenas na detecção, mas também na oferta de soluções pertinentes para desafios de segurança. A abordagem proposta não apenas reforça a importância da segurança da informação no contexto da ISO/IEC 27002:2013 e da cibersegurança, mas também ressalta a contribuição significativa da IA para fortalecer a resiliência contra ameaças digitais.

Com a popularização da IA Generativa e sua acessibilidade para os desenvolvedores, este trabalho demonstrou uma aplicação prática do potencial transformador da IA aplicada a Segurança de Sistemas de Informação. Dessa forma, este trabalho apresenta um possível caminho para a redução dos motivos de não implementação de medidas de segurança, conforme apontados por Costa et. al, 2018 (falta de conhecimento e prazos apertados).

Trabalhos futuros irão explorar integração com outras plataformas e realizar uma pesquisa de aceitação de tecnologia com desenvolvedores.

REFERENCES

[1] ABNT NBR ISO/IEC. *Tecnologia da informação - técnicas de segurança - código de prática para controles de segurança da informação.* Number 27002 in 27000. ABNT, October 2013. ISBN 9788507046134.

[2] Trend Micro. Stepping ahead of risk: Trend Micro 2023 Midyear Cybersecurity Threat Report. Technical report, 2023. URL https://documents.trendmicro.com/images/TEx/articles/Risk_Landscape_infographic-aypd962.png.

[3] Raquel Hengen Ribeiro. *Identificação de bugs em código-fonte usando aprendizagem de máquina.* Trabalho de Conclusão de Curso (Ciência da Computação), Universidade Federal da Fronteira Sul - UFFS, September 2021. URL https://rd.uffs.edu.br/handle/prefix/5000.

[4] Pablo V. Costa, Willian I. Gonçalves, Eliezer Dutra Gonçalves, and Nilson M. Lazarin. Nível de conhecimento de desenvolvedores sobre segurança em aplicações web: Pesquisa e análise. In *Anais da V Escola Regional de Sistemas de Informação do Rio de Janeiro*, pages 92–99, Porto Alegre, RS, Brasil, 2018. SBC. URL https://doi.org/10.5753/ersirj.2018.4661.

[5] Matheus Cristani, Waldenir Alves, Gabriel Pereira, and Nilson Lazarin. Um breve panorama sobre a disciplina de segurança nos cursos de sistemas de informação no brasil. In *Anais Estendidos do XVI Simpósio Brasileiro de Sistemas de Informação*, pages 1–4, Porto Alegre, RS, Brasil, 2020. SBC. URL https://doi.org/10.5753/sbsi.2020.13114.

[6] Samir Ibrahim Elias. O impacto da inteligÊncia artificial no comportamento organizacional. *Revista Ilustração*, 4(3):33–39, set. 2023. URL https://doi.org/10.46550/ilustracao.v4i3.176.

[7] Alex Bruno Gonçalves Almeida and Juliana Lilis da Silva. Inteligência artificial aplicada à segurança da informação. *Perquirere*, 2(17):241–254, 2020. URL https://revistas.unipam.edu.br/index.php/perquirere/article/view/2040.

[8] Erik Henrique Leite and Douglas Francisco Ribeiro. O papel transformador da inteligÊncia artificial na seguranÇa. *Revista Interface Tecnológica*, 20(1):181–190, jun. 2023. URL https://doi.org/10.31510/infa.v20i1.1669.

[9] Pierre Carbonnelle. Top IDE index, 2023. URL https://pypl.github.io/IDE.html.

[10] Luciano Rossoni and Chat Gpt. A inteligência artificial e eu: escrevendo o editorial juntamente com o ChatGPT. *Revista Eletrônica de Ciência Administrativa*, 21(3):399–405, October 2022. URL https://doi.org/10.21529/RECADM.2022ed3.

[1] https://marketplace.visualstudio.com/items/MustafaNeto.mn-analise/
[2] https://youtu.be/2G4MxDEMqUg
[3] https://github.com/snoopysecurity/Vulnerable-Code-Snippets