Security testing of information systems using artificial intelligence

V. V. Vaskevich, M. A. Vinokurov, T. A. Melnikova, D. A. Shumilo
Francisk Skorina Gomel State University, Gomel, Belarus,

Abstract*. This paper provides an overview of the main functions of popular generative artificial intelligence chatbots: ChatGPT, Grok, DeepSeek, Gemini, and Copilot. It presents the results of using artificial intelligence as an assistant for system vulnerability testing and the analysis of information security logs. The main advantages and disadvantages of using AI chatbots in the field of cybersecurity are described.*

I. Introduction

In the digital era, cybersecurity has become particularly crucial. Artificial intelligence (hereinafter referred to as AI) and its rapid development are introducing significant changes to the methods of penetration testing and vulnerability analysis of modern systems. These systems are constantly growing in complexity, which requires more comprehensive approaches to ensuring their security.

This paper reviews popular generative artificial intelligence chatbots and describes scenarios for their use as a tool to assist specialists in penetration testing and log analysis in modern security systems.

II. Overview of popular generative AI chatbots

From an accessibility standpoint, generative AI chatbots are optimal solutions for assisting novice specialists [1-2].

One of the very first chatbots was ChatGPT from OpenAI, with its first version released in 2022. Without registration, the bot offers an introductory set of features and a limited number of queries. It processes only text and program code, providing answers based on its internal knowledge base or web searches. Upon registration, the bot provides a voice assistant feature that simulates emotions, as well as a query history and the ability to return to previous conversations. Its capabilities expand to reading uploaded files, images, and audio. With a paid subscription, ChatGPT offers the same functions but improved and without limits. Currently, OpenAI blocks access to its services for users from Russia, Belarus, China, Afghanistan, North Korea, and many other countries.

Grok is a chatbot developed by xAI in 2023. Without registration, Grok provides users with answers based on its knowledge base, web searches, and posts on the social network X. The number of queries is limited to four. For registered users, Grok expands its functionality, offering access to "DeepSearch" and "Think Mode," where the model broadens its search and outlines its reasoning. The number of possible queries increases, but images and files are still accepted in limited quantities. It maintains a chat history. A paid subscription, SuperGrok, removes query limits and provides enhanced functionality. Unlike ChatGPT, despite being unavailable in some countries, Grok is open for registration to everyone.

DeepSeek is a chatbot developed by the eponymous Chinese startup in 2023. It is completely free and available for CIS countries, though registration is required for use. Due to its budget-friendly nature, its functionality is quite limited. The "DeepThink" function demonstrates the bot's reasoning, while "Search" is responsible for searching and analyzing information strictly within uploaded files (PDF, DOCX, XLSX, PPTX, and TXT) or text. The chat lacks the ability to browse internet resources in real-time, and its knowledge base is limited to information up to July 2024.

Gemini is a chatbot from Google. The first version was released under the name Bard in February 2023, and a year later, Gemini 1.5 appeared. Unregistered users have a wide range of interaction possibilities with this chatbot [3]. The "DeepResearch" function allows Gemini to conduct a more in-depth information search. Additionally, Gemini offers a range of integrations with Google services. The free version has query limits, while paid subscriptions increase this number in proportion to the subscription cost. Gemini is unavailable in some regions, including Belarus and Russia.

Copilot is a chatbot from Microsoft, developed in 2023. It is a multilingual model. Without registration, the bot can hold a conversation, analyze and generate images, and provides access to the "Think Deeper" function. Upon registration, Copilot saves chat history. A much broader range of possibilities opens up with a Copilot Pro subscription: the chat can be used within Microsoft applications and provides better-generated responses. The free version of the model is available in CIS countries; however, purchasing a subscription is not possible due to restrictions from Microsoft.

Despite the fact that most of the listed chatbots are unavailable in Belarus and Russia, many users access them using bypass mechanisms.

III. Using AI in penetration testing

Penetration testing remains one of the key methods for assessing security posture.

The initial stage of any security audit involves gathering as much information as possible about the target object. Gemini (version 2.5 pro from Google AI Studio) was employed to process large volumes of open-source intelligence (OSINT) and assisted in the automated search and structuring of information about domain names, associated IP addresses, and other data based on Nmap scan results. The advantage of using Gemini at this stage was its ability to quickly process and aggregate information from multiple sources, as well as highlight potentially interesting facts that might be overlooked during a manual search. For example, Gemini was used to generate specific search queries to detect vulnerable configurations.

After the initial information gathering, the next step was scanning for known vulnerabilities. Here, Gemini was applied to analyze the results obtained from standard scanners. It helped interpret their reports, explain the nature of the identified vulnerabilities, and provided information on possible exploitation methods based on its knowledge base. Gemini was also used to correlate information about the software versions used in the system with public vulnerability databases (CVE). This allowed for a faster identification of critical points requiring special attention and helped generate examples of potentially vulnerable configuration files and code snippets for subsequent manual verification.

It is important to note that Gemini was not used for the direct automated exploitation of vulnerabilities, as this requires deep integration into the toolchain and cannot be accomplished without significant resource investment. However, the model assisted in the theoretical development of attack vectors. For instance, based on a description of a specific vulnerability and system configuration, Gemini proposed possible sequences of actions for its verification.

In the final stage, during the preparation of the test report, Gemini was used to structure information and draft preliminary sections. It helped summarize the found vulnerabilities, suggested recommendations for their remediation, and presented technical information in a more comprehensible format. This did not replace expert judgment and final editing but significantly accelerated the documentation preparation process.

IV. Using AI in log analysis

One of the most labor-intensive tasks for a modern cybersecurity analyst is the analysis of logs and artifacts left by malware. System journals, network traffic, registry changes – all of this constitutes a massive amount of data, the manual interpretation of which requires considerable time and expertise. However, generative artificial intelligence, as exemplified by the use of ChatGPT-4, opens new possibilities for automating this process, acting as a powerful assistant. This is most evident when working with logs generated in isolated environments, such as the ANY.RUN sandbox. After analyzing a suspicious file, the system provides a detailed report of its actions. Instead of manually correlating events, ChatGPT-4 was used to obtain a ready-made summary. When analyzing an executable file, ChatGPT-4 instantly highlighted key suspicious activities recorded in the logs: execution from a user's temporary folder, file creation, and establishing connections to non-standard network ports. Based on this disparate data, a conclusion was formed regarding the type of threat corresponding to the events in the logs.

The understanding of malicious activity thus obtained flows smoothly into the next task: creating mechanisms for future threat detection. Here, the very same artifacts identified in the logs become the basis

for generating detection rules. Based on network traffic logs, ChatGPT-4 can assist a specialist in formulating rules for the Suricata system. The rules proposed by ChatGPT-4 are not always perfect and require minor adjustments by a specialist–for instance, specifying string encoding or correcting logic. Nevertheless, they serve as an excellent starting point, significantly speeding up the process [4].

V. Conclusion

The experience of using AI in penetration testing and security log analysis has demonstrated its potential as an auxiliary tool. It can not only accelerate the collection and analysis of information but also aid in understanding vulnerabilities and preparing the groundwork for reporting. In some cases, AI can assist in writing or adapting simple scripts to automate routine checks, such as brute-forcing standard credentials on available services or checking the availability of specific ports and services. This allows a specialist to save time on routine tasks and focus on more complex aspects.

Nevertheless, it is crucial to understand that AI lacks the intuition and real-world experience of a security specialist. It operates on the data it was trained on and cannot always adequately assess the specific context of a particular system. The critical thinking and expert judgment of a specialist remain indispensable at all stages of testing.

Thus, AI does not replace the analyst but rather acts as an effective tool that takes on the routine work of interpreting logs and creating signatures, allowing the human to focus on more complex aspects of the investigation and respond to threats more quickly.

The use of AI should be viewed as a means of enhancing human efficiency, not as a complete replacement. The further development of such tools will make them even more valuable in the arsenal of cybersecurity specialists.

References

[1] *A. Saggu, L. Ante* “The influence of ChatGPT on artificial intelligence related crypto assets: Evidence from a synthetic control analysis Author links open overlay panel”, Finance Research Letters Vol.55(B), 2023,

103993 (https://doi.org/10.1016/j.frl.2023.103993)

[2] *Т.А. Безуглый, М.Е. Ершова,* “Использование текстовых нейросетей и искусственного интеллекта в учебных работах студентов”, Проблемы современного образования, 2023. pp. 206–216.

[3] С. Башкиров “Gemini 2.5 Pro: что умеет самая мощная ИИ-модель от Google” [Electronic resource] Access mode: https://trends.rbc.ru/trends/industry/67ea6cc99a794758d63b84e8.

[4] *ANY.RUN* “How to Use ChatGPT for Malware Analysis” [Electronic resource] Access mode: https://hackernoon.com/how-to-use-chatgpt-for-malware-analysis.