CyberAI: A Live Demonstration of Wi-Fi Threats and AI-Powered Defenses

Paper #123

Abstract. Public Wi-Fi networks have become a common thing in modern life, however, they are still a source of danger in the digital field. Attacks like Rogue Access Point or Man in The Middle allow malicious actors to create fake access points and intercept user traffic. The project "CyberAI" presents an integrated system with two modules: Rogue AP for the demonstration of the attacks and AI-scanner for the detection of potentially dangerous networks. Rogue AP works on the basis of Raspberry Pi, collecting metadata of the connected devices. AI-scanner using a Wi-Fi adapter in monitor mode, collecting network characteristics, and sending them to GPT API, which rates the level of suspicion. However, a key innovation of our system is its explainability - AI gives us human-readable justification as to why each network got exactly this score. During the demonstration, participants will see a full cycle of an attack and defense in real-time. The project has the purpose of increasing awareness about cyber hygiene and demonstrating the potential of AI in the context of analysis of cybersecurity.

1 Introduction

Nowadays, public Wi-Fi networks are an essential part of digital infrastructure - they are used in cafes, airports, and universities. However, people usually connect to them without any doubts, without any checking, and without realizing the potential risks. Lack of awareness about the threats associated with such networks makes them attractive targets for attackers.

One of the simplest and at the same time most effective attacks is a Rogue Access Point - creating a fake WiFi spot, that imitates a legitimate network. When the user connects to such a network, a threat actor can intercept traffic, collect DNS/SNI requests, and even perform phishing via a captive portal or replace the websites that the user visits. This type of attack requires minimal technical resources but has serious consequences.

In this context, the CyberAI project offers a demonstration of both the attack and the means of detecting it. Existing approaches to threat detection are typically based on signature-based intrusion detection systems (IDS) or simple heuristic rules (e.g., identifying open networks or weak encryption types). Such methods often produce a large number of false positives or are unable to detect new, previously unknown attacks. Some commercial solutions use machine learning algorithms, but mostly in the form of “black boxes” without a clear explanation of the results. In contrast, our system uses a large language model (LLM) for contextual analysis based solely on metadata. Importantly, the model not only determines the level of suspicion, but also provides an explanation of why this particular network is considered threatening. The purpose of this demonstration is not only to visualize the vulnerability, but also to offer a transparent and intuitive method of threat detection based on AI. In this article, we describe the architecture of the CyberAI system and the demonstration scenario.

2 System Architecture

2.1 The Rogue Access Point Module

The Rogue AP module [2] is implemented on the basis of a Raspberry Pi 4B microcomputer [3] using the hostapd (access point creation) and dnsmasq (DHCP and DNS server) utilities. The device emulates a public network with a trustworthy name (e.g., “Starbucks_WiFi”). After connecting, the user gets access to the Internet. After connecting, the device logs MAC addresses, hostnames, as well as DNS and SNI requests to monitor the domains visited. This data allows for OSINT analysis and private espionage operations.

2.2 The AI-Powered Scanner Module

The security module consists of a Raspberry Pi 4 B with a Wi-Fi adapter in monitoring mode, which collects information about all available networks using airodump-ng [1]. A Python script parses the output CSV files and forms a structured set of characteristics for each network: BSSID, SSID, vendor (based on MAC address), signal strength (RSSI), encryption type, number of clients, as well as contextual features of the network name (e.g., use of keywords or symbols).

This metadata is sent to the GPT 4 API [4] as a prompt. The model analyzes the data, returns a suspicion score (0–10), and a text explanation of why this particular network is considered a threat. The response is parsed by a JSON interpreter and then displayed to the user.

3 The Demonstration Scenario

Step 1: Activation of a rogue access point in a public environment. A network with an attractive name (imitating a typical Wi-Fi name in a café)

Step 2: The victim device (e.g., a smartphone) connects to the open fake network.

Step 3: The attacker’s screen displays connection metadata: MAC addresses, hostnames, and DNS/SNI requests.

Step 4: The AI scanner is launched on another laptop. It scans all surrounding networks, analyzes them, detects the fake hotspot, and assigns a high level of suspicion. The visitor sees not only the result, but also an explanation from the AI justifying its decision.

4 Innovation and Contribution

Unlike traditional IDS/IPS systems based on signatures or simple heuristics, the AI module uses a large language model for contextual analysis of network metadata. This allows the detection of more complex or new attack vectors that do not have previous patterns.

The main advantage is the explainability of the result. The AI model not only rates suspicious activity, but also explains the reason for its decision, which is important for both analysts and ordinary users. This approach paves the way for the creation of a new generation of systems capable of autonomous but transparent decision-making in the field of cyber security.

5 Conclusion

The CyberAI project demonstrates both attacks and innovative defenses in the field of wireless security. The combination of practical demonstration and AI analytics allows not only to identify the threat, but also to explain its essence. This helps raise awareness, demonstrates the potential of large language models in cyber defense, and lays the foundation for the creation of intelligent systems with a high level of transparency.

Ultimately, such an interactive demonstration is an effective means of communicating complex technical issues to a wide audience.

6 Citations and references

References

[1] Aircrack-ng Team. Aircrack-ng. https://www.aircrack-ng.org/, 2025.

[2] J. Bellardo and S. Savage. 802.11–denial-of-service attacks: real vulnerabilities and practical solutions. In *Proceedings of the 12th USENIX security symposium*, pages 15–28, 2003.

[3] G. Foser, E. Pimenidis, C. Kouroupetroglou, and P. Polydoros. Raspberry pi as a wireless attack platform. In *2015 4th International Conference on Advanced Information Technologies and Applications (ICAITA)*, pages 50–55. IEEE, 2015.

[4] OpenAI. Gpt-4 technical report. Technical Report arXiv:2303.08774, OpenAI, 2023. URL https://arxiv.org/abs/2303.08774.

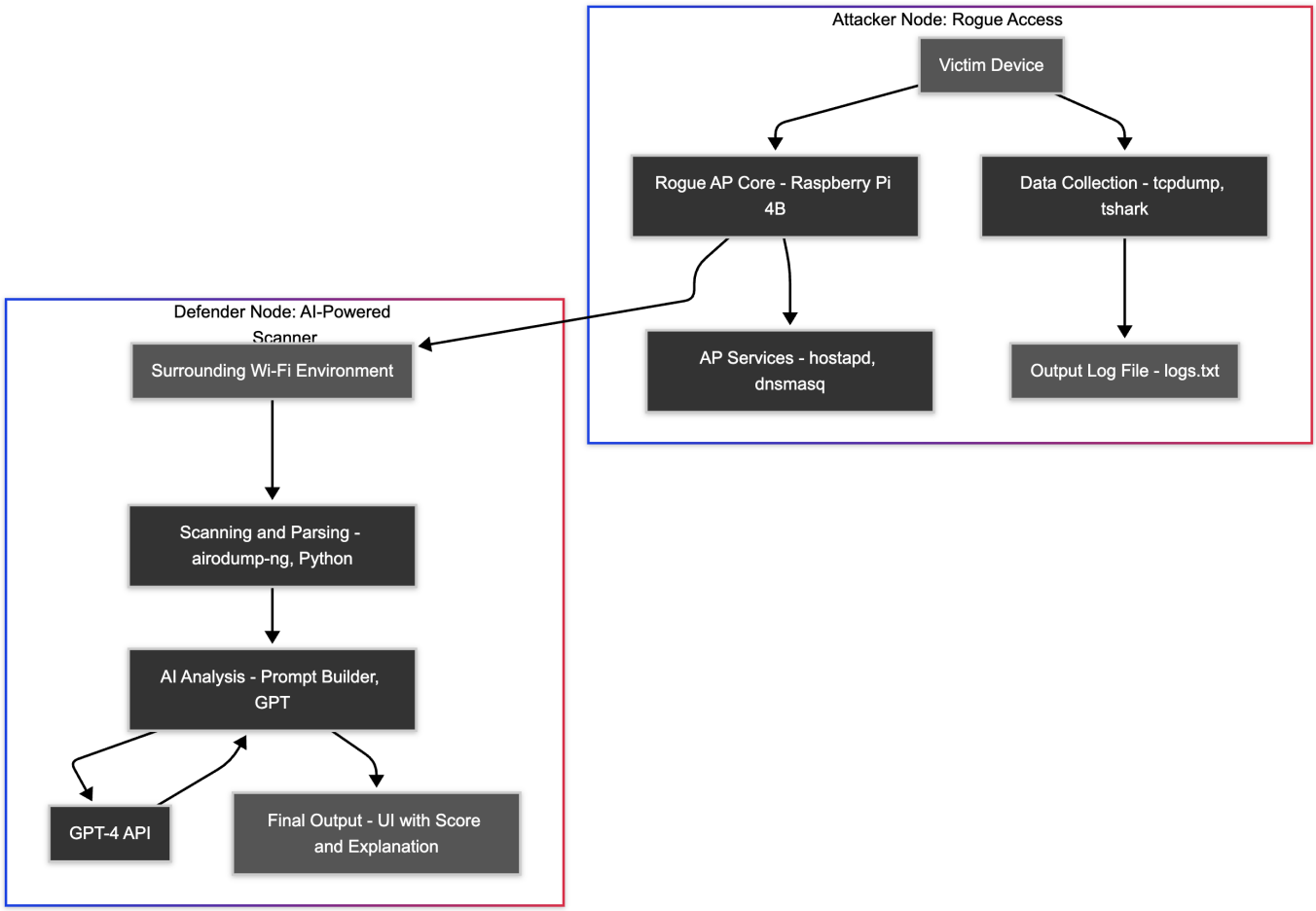

Figure 1. The architecture of the CyberAI system, which consists of two main parts. The top part of the diagram illustrates the Attacker Node, which simulates a Wi-Fi attack. The bottom part shows the Defender Node, which detects the threat using an AI model.