World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:18, No:05, 2024

The Road Ahead: Merging Human Cyber Security Expertise with Generative AI

Authors : Brennan Lodge

Abstract : Cybersecurity professionals have long been embroiled in a digital arms race, confronting increasingly sophisticated threats with innovative solutions. The field of cybersecurity is in an unending race against malicious adversaries. As threats evolve in complexity, the tools used to defend against them need to advance even faster. Burdened with a vast arsenal of tools and an expansive scope of threat intelligence, analysts frequently navigate a complex web, trying to discern patterns amidst information overload. Herein lies the potential of Retrieval Augmented Generation (RAG). By combining the capabilities of Large Language Models (LLMs) with a generative AI facet, RAG brings to the table an unparalleled ability for real-time cross-referencing, bridging the gap between raw data and actionable insights. Imagine an analyst named Sarah working at a global Fortune 500 company. Every day, Sarah navigates a maze of diverse knowledge bases, real-time threat intelligence, and her company's vast proprietary data, from network specifics to intricate technical blueprints. One day, she's challenged by a potential breach through a personal device due to the company's global "Bring Your Own Device" policy. With the clock ticking, Sarah has mere minutes to trace the malware's origin, all while considering complex regional regulations. As she races against the benchmark of Mean Time To Resolution (MTTR), she wonders: Could "Cozy Bear" with its notorious malware tactic, HAMMERTOSS, be behind this? Balancing policy intricacies, global network considerations, and ever-emerging cyber threats, Sarah's role epitomizes the intense challenges faced by today's cybersecurity analysts. While analysts grapple with this array of intricate, time-sensitive challenges, the necessity for precision and efficiency is key. RAG technology—a cutting-edge advancement in Gen AI—is a promising solution. Designed to assimilate diverse data sources such as cyber advisory notices, phishing email sentiment, secure and insecure code examples, information security policy documentation, and the MITRE ATT&CK framework, RAG equips analysts with real-time querying capabilities through a vector database and a cross referenced concise response from a Gen AI model. Traditional relational databases often necessitate a tedious process of filtering through numerous entries. Now, with the synergy of vector databases and Gen AI models, analysts can rapidly access both contextually or semantically akin data points. This augmented approach equips analysts with a comprehensive understanding of the prevailing cyber threats, elevating the robustness of cybersecurity defenses and upskilling the analyst and team, too. Vector databases underpin the knowledge translation in Gen AI. They bridge the gap between raw data and translation into meaningful insights, ensuring that analysts are equipped with comprehensive and relevant information. This superior capability of the RAG framework, with its impressive depth and precision, finds application across a broad spectrum of cybersecurity challenges. Let's delve into some use cases where its potential becomes particularly evident: Phishing Email Sentiment Analysis: Phishing remains a predominant vector for cybersecurity breaches. Leveraging RAG's capabilities, analysts can not only assess the potential malevolence of an email but can also understand the context behind it. By cross-referencing patterns from varied data sources in real-time, the detection process evolves from a mere content evaluation to a holistic understanding of attacker tactics, behaviors, and evolving profiles. This allows for the identification of nuanced phishing strategies that might otherwise go undetected. Insecure Code Analysis: Software vulnerabilities form a critical entry point for cyber adversaries. With RAG, the process of code evaluation undergoes a transformation. Instead of manual code reviews, the system pulls insights from vector databases and historical code snippets marked as insecure, enabling detection of vulnerabilities based on historical patterns, emerging threat vectors, and even predictive threat modeling. This ensures that even the most obfuscated or embedded vulnerabilities are identified, and corrective measures can be promptly implemented. Vulnerability and Upskill Advisory: In the fast-paced world of cybersecurity, staying updated is paramount. Through RAG's capabilities, analysts are not only made aware of real-time vulnerabilities but are also guided on the necessary skills and tools needed to combat them. By dynamically sourcing data through vulnerability advisories, news on advanced persistent threats, and tactics to defend, RAG ensures that analysts are not only reactive to threats but are also proactively upskilled, thereby bolstering their defense mechanisms. Information Security Policies for Compliance Teams: Compliance remains at the heart of many organizational cybersecurity strategies. However, with ever-shifting regulatory landscapes, staying compliant becomes a moving target. RAG's ability to source real-time data ensures that compliance teams always have access to the latest policy changes, guidelines, and best practices. This not only facilitates adherence to current standards but also anticipates future shifts, assists with audits, and ensures that organizations remain ahead of the compliance curve. Fusing a RAG architecture with platforms like Slack amplifies its practical utility. Slack, known for its real-time communication prowess, seamlessly evolves into more than just a messaging platform in this context. Cybersecurity analysts can pose intricate queries within Slack and, almost instantaneously, receive comprehensive feedback powered by the harmonious interplay of RAG and Gen AI. This integration effectively transforms Slack into an AI-augmented chatbot-like assistant for cybersecurity professionals, always ready to provide informed insights on-demand, making it an indispensable ally in the ever-evolving cyber battlefield. Navigating the vast landscape of cybersecurity, analysts often encounter unfamiliar terminologies and techniques., analysts require tools that not only detect or inform them of threats, like CISA (U.S Cybersecurity Infrastructure Security Agency) Advisories, but also interpret and communicate them effectively. Consider a junior cybersecurity analyst named Alex, who comes across the term "Kerberoasting" while reviewing a network log. Unfamiliar with its intricacies, Alex turns to Slack to

pose a query: "chat explain is Kerberoasting, using CISA." Almost instantaneously, Slack, powered by the harmonious interplay of RAG and Gen AI, provides a detailed response, cross-referencing a recent cyber advisory on the technique. It explains how attackers can exploit the Kerberos Ticket Granting Service to decipher service account passwords, potentially compromising a network. In this dynamic realm of cybersecurity, the blend of RAG and Generative AI represents more than just a technological leap. It embodies a paradigm shift, promising a future where human expertise and AI-driven precision join forces. As cyber threats continue their relentless advance, this synergy ensures that defenders are equipped with an arsenal that's not just reactive, but also profoundly insightful. No longer should analysts be submerged in a deluge of data without direction. Instead, they should be empowered, to discern, act, and preempt with unparalleled clarity and confidence. By harmoniously intertwining human discernment with AI capabilities, we should chart a path towards a future where cybersecurity is not just about defense, but about achieving a strategic advantage, paving the way for a safer, informed and a more secure digital horizon.

Keywords : cybersecurity, gen AI, retrieval augmented generation, cybersecurity defense strategies
Conference Title : ICCSAI 2024 : International Conference on Cyber Security and Artificial Intelligence
Conference Location : Bucharest, Romania
Conference Dates : May 16-17, 2024