Wednesday, August 19

12:10 PM - 12:30 PM

Hardening networks against cryptographic threats

Mark Maglin

VP Cybersecurity

ECS Federal

Abstract:

"Developing validated, resilient Machine-to-Machine (M2M) security protocols to prevent ""Synthetic Access"" breaches driven by autonomous agentic AI is crucial, especially given that current protocols like the Model Context Protocol (MCP) often offer considerable freedom in design, leading to ambiguous and potentially insecure usage. This flexibility creates new attack paths, as servers might query and execute actions for connected clients, inverting traditional interaction patterns and making them vulnerable to arbitrary code execution (ACE), privilege escalation, and data exfiltration by autonomous agents.

A multi-layered, adaptive strategy is required. This centers on adopting Secure-by-Design frameworks like Google's Secure AI Framework (SAIF) to embed lifecycle-wide security. It must feature robust identity and access control, using least-privilege and zero-trust principles to grant AI agents separate identities and fine-grained permissions. Continuous monitoring and detection are essential, utilizing real-time behavioral monitoring, chain-of-thought analysis, and anomaly detection alongside AI ""supervisors"" to identify misalignment, unexpected actions, or injection attempts while suppressing evasion. Additionally, architectural resilience and isolation via clear trust boundaries, environment isolation, and sandboxing will limit blast radiuses and prevent cascading failures.

Ensuring secure communication requires message authenticity, confidentiality, and replay protection through cryptographic signatures and strict validation. Proactive prevention and response strategies must combine automated blocking, human-in-the-loop interventions for high-risk scenarios, and instance shutdown capabilities. The integrity of the agentic supply chain must be reinforced through provenance tracking, verification, and continuous vulnerability management for all tools and components.

Finally, adversarial testing and continuous feedback loops, such as regular red-teaming, are critical to validate resilience and refine evolving protocols. Comprehensive auditing and logging remain foundational for incident response and accountability.

Attendees will learn or gain:

A deep understanding of novel AI agent threats: Insights into how autonomous agentic AI introduces new attack vectors like agent goal hijack, tool misuse, and memory poisoning, specifically targeting M2M protocols.

Practical strategies for architectural and operational security: Knowledge of frameworks (e.g., SAIF, AI Control Roadmap), core security principles (e.g., zero-trust, least privilege), and technical controls (e.g., agent identity, sandboxing, secure supply chain) to build resilient M2M protocols, principles (e.g., zero-trust, least privilege), and technical controls (e.g., agent identity, sandboxing, secure supply chain) to build resilient M2M protocols.

Methods for continuous validation and adaptive defense: Approaches to employing adversarial testing, comprehensive logging, and real-time behavioral monitoring to ensure M2M security protocols can withstand and adapt to evolving AI-driven threats.