AL2025_44 AI-Driven Phishing Campaign Using LLM-Crafted SVG Files (September 30th, 2025)

Description

Microsoft has identified a new phishing campaign targeting organizations that leverages large language models (LLMs) to create obfuscated Scalable Vector Graphics (SVG) files. These files are used to bypass traditional email security defenses by mimicking legitimate business-related content. The campaign highlights the growing adoption of AI-driven obfuscation techniques by threat actors to create more convincing phishing lures and evade detection.

Attack Details

In this campaign, attackers gained initial access by using compromised business email accounts to send phishing messages disguised as file-sharing notifications. Instead of delivering a PDF, the emails contained malicious SVG files that appeared harmless but were embedded with obfuscated scripts. The SVG code was crafted to mimic a legitimate business analytics dashboard and was filled with verbose, business-related terminology such as “revenue,” “growth,” “quarterly,” and “operations” to disguise its true intent. When executed, the SVG redirected victims to a CAPTCHA page, which then led to a fake login portal designed to harvest credentials. Microsoft Threat Intelligence concluded that the code was likely generated with the aid of a large language model (LLM), citing indicators such as overly descriptive variable and function names, verbose comments, redundant structures, and a modular, over-engineered design not typical of manually written code.

Remediation

- **Email Security Enhancements:**
    - Block or flag unexpected SVG attachments in corporate environments
    - Strengthen filtering for self-addressed emails with hidden BCC recipients
- **User Awareness:**
    - Train staff to be cautious with file-sharing notifications, especially when the file type differs from expectations (e.g., PDF vs. SVG)
- **Technical Controls:**

  - Enable advanced attachment scanning that inspects text-based file formats like SVG for embedded scripts
  - Deploy behavioral detection that identifies redirect chains and credential-harvesting sites

- **Incident Response:**
  - Monitor for unauthorized login attempts or credential use
  - Reset credentials for any users who interacted with suspicious SVG attachments

The Guyana National CIRT recommends that users and administrators review these updates and apply them where necessary.

References

- Microsoft Threat Intelligence Advisory. (September 29, 2025). Retrieved from The Hacker News. https://thehackernews.com/2025/09/microsoft-flags-ai-driven-phishing-llm.html

- Microsoft Threat Intelligence Advisory. (September 24, 2025). Retrieved from Microsoft Threat Intelligence. https://www.microsoft.com/en-us/security/blog/2025/09/24/ai-vs-ai-detecting-an-ai-obfuscated-phishing-campaign/?msockid=314f2b75bc1067d701cd3f71bdae6663