精选快讯　软件快讯　2026年11月13日 13:35

研究人员演示利用窃取凭证将Claude Code转为远程工具窃取1GB敏感文件

要点

- 网络安全研究人员在受控实验中演示了一种利用窃取API凭证将Anthropic的Claude Code转变为远程受控工具的攻击方法。
- 攻击的前提条件是犯罪分子首先入侵开发人员的计算机并获取有效的Claude凭证。
- 在获得对运行Claude Code的Linux计算机的访问权限后，AI代理执行了九项连续任务。

网络安全研究人员在受控实验中演示了一种利用窃取API凭证将Anthropic的Claude Code转变为远程受控工具的攻击方法。该攻击旨在窃取敏感企业信息，实验表明攻击者无需发现新的软件漏洞或绕过平台安全保护，即可利用现有AI能力实施入侵。研究人员强调，该技术目前仅在实验室环境中演示，尚未在现实世界的攻击中被观察到，且仅在Claude的记忆功能上进行了测试，尚未在其他AI平台上复现。

攻击的前提条件是犯罪分子首先入侵开发人员的计算机并获取有效的Claude凭证。研究人员发现，Claude的记忆存储（memory stores）可以通过API访问，包括在未使用Claude Managed Agents的工作区中，这使其成为外部操作员与AI代理之间的通信渠道。在模拟的企业网络环境中，实验使用了微软的Active Directory环境，以验证攻击在典型企业架构中的可行性。

在获得对运行Claude Code的Linux计算机的访问权限后，AI代理执行了九项连续任务。代理首先检查本地网络，收集员工账户信息，并访问了包含用户名和密码详情的匿名FTP共享。随后，代理生成了190个潜在用户名，识别了与域管理员账户关联的密码，并定位了可访问的网络资源。最终，代理下载了一个1 GB的文件，并通过Anthropic的Files API分三次传输了这些信息，完成了数据窃取过程。

报道来源：the420

来源网站：https://www.inews.today/