วันที่ 26 กุมภาพันธ์ 2569

ช่องโหว่ร้ายแรงใน Claude Code เสี่ยงถูกสั่งรันโค้ด-ขโมย API Key

นักวิจัยด้านความปลอดภัยไซเบอร์ตรวจพบช่องโหว่หลายจุดใน Claude Code (ผู้ช่วยเขียนโค้ดพลัง AI จากค่าย Anthropic) ซึ่งช่วยให้แฮกเกอร์สามารถสั่งรันโค้ดจากระยะไกล (Remote Code Execution) และขโมยรหัสเข้าถึงบริการ (API Credentials) ของผู้ใช้งานได้

รายงานจาก Check Point Research ที่ส่งถึง The Hacker News ระบุว่า "ช่องโหว่เหล่านี้อาศัยกลไกการตั้งค่าหลายรูปแบบ เช่น Hooks, เซิร์ฟเวอร์ Model Context Protocol (MCP) และตัวแปรสภาพแวดล้อม (Environment Variables) เพื่อรันคำสั่งเชลล์ (Shell Command) และดึง Anthropic API Key ออกไป เพียงแค่ผู้ใช้ดาวน์โหลด (Clone) และเปิดโปรเจกต์ที่ไม่น่าเชื่อถือเท่านั้น"

เจาะลึก 3 ช่องโหว่หลักที่ถูกค้นพบ นักวิจัยแบ่งความเสี่ยงออกเป็น 3 กลุ่ม โดยเรียงตามลำดับเวลาและการแก้ไข ดังนี้

1) **No CVE (CVSS 8.7)**

   รายละเอียดและผลกระทบ: ช่องโหว่แบบ Code Injection แฮกเกอร์สามารถข้ามขั้นตอนขอความยินยอมจากผู้ใช้และสั่งรันโค้ดผ่านไฟล์ .claude/settings.json ได้ทันทีเมื่อเปิดโฟลเดอร์ใหม่

   สถานะแก้ไข: เวอร์ชัน 1.0.87 (ก.ย. 2025)

2) **CVE-2025-59536 (CVSS 8.7)**

   รายละเอียดและผลกระทบ: ช่องโหว่ Code Injection รันคำสั่งเชลล์อัตโนมัติเมื่อเริ่มต้นเครื่องมือ หากผู้ใช้เปิด Claude Code ในโฟลเดอร์ที่ไม่น่าเชื่อถือ

   สถานะแก้ไข: เวอร์ชัน 1.0.111 (ต.ค. 2025)

3) **CVE-2026-21852 (CVSS 5.3)**

   รายละเอียดและผลกระทบ: ช่องโหว่การเปิดเผยข้อมูลในขั้นตอนโหลดโปรเจกต์ของ Claude Code ที่เปิดทางให้รีโพซิทอรีอันตรายดึงข้อมูลออกไปได้ รวมถึง Anthropic API key

   สถานะแก้ไข: เวอร์ชัน 2.0.65 (ม.ค. 2026)

กลไกการโจมตี: แค่ “เปิดโฟลเดอร์” ก็เสี่ยงแล้ว

ในกรณีของ CVE-2026-21852 Anthropic อธิบายว่า หากแฮกเกอร์สร้างโปรเจกต์ที่กำหนดค่า ANTHROPIC_BASE_URL ให้ชี้ไปยังเซิร์ฟเวอร์ของตนเอง Claude Code จะส่งคำขอ API ไปยังที่อยู่นั้นทันที ทำให้ API Key ของนักพัฒนารั่วไหลได้ ก่อนที่หน้าต่างแจ้งเตือน (Trust Prompt) จะปรากฏขึ้นด้วยซ้ำ

ผลกระทบที่อาจเกิดขึ้น:

หากผู้โจมตีใช้ช่องโหว่เหล่านี้สำเร็จ อาจทำให้พวกเขาสามารถ

- เข้าถึงไฟล์โปรเจกต์ที่แชร์อยู่
- แก้ไข ลบ หรืออัปโหลดข้อมูลอันตรายไปยังระบบคลาวด์
- ดำเนินการผ่าน API โดยที่เจ้าของบัญชีไม่รู้ตัว
- ทำให้เกิดค่าใช้จ่าย API โดยไม่คาดคิด

นิยามใหม่ของความเสี่ยงในยุค AI

สำหรับช่องโหว่ CVE-2025-59536 มีเป้าหมายที่คล้ายคลึงกัน โดยความแตกต่างหลักคือ ผู้โจมตีจะใช้ประโยชน์จากการตั้งค่าที่กำหนดไว้ในรีโพซิทอรีผ่านไฟล์ .mcp.json และ .claude/settings.json เพื่อข้ามขั้นตอนการขออนุมัติจากผู้ใช้ก่อนที่จะโต้ตอบกับเครื่องมือหรือบริการภายนอกผ่าน Model Context Protocol (MCP) โดยสามารถทำได้เพียงแค่กำหนดค่า "enableAllProjectMcpServers" ให้เป็น true

บทสรุปจากนักวิจัย:

Check Point Research ระบุว่า เมื่อเครื่องมือ AI สามารถรันคำสั่ง เชื่อมต่อบริการภายนอก และสื่อสารเครือข่ายได้เอง ไฟล์ตั้งค่าจะไม่ใช่เพียงแค่ “บริบท” อีกต่อไป แต่เป็นส่วนหนึ่งของชั้นการทำงานจริงของระบบ

“ความเสี่ยงไม่ได้จำกัดอยู่ที่การรันโค้ดจากแหล่งไม่น่าเชื่อถือ แต่ขยายไปถึงการ เปิดโปรเจกต์ที่ไม่น่าเชื่อถือ ในสภาพแวดล้อมการพัฒนาแบบ AI ห่วงโซ่อุปทานของซอฟต์แวร์ไม่ได้เริ่มที่ซอร์สโค้ดอีกแล้ว แต่รวมถึงระบบอัตโนมัติที่รายล้อมด้วย”

ข้อมูลอ้างอิง

Feb 25, 2026, By Ravie Lakshmanan

- https://thehackernews.com/2026/02/claude-code-flaws-allow-remote-code.html