

Ransomware affiliates ran intrusions through AI agents, an exposed server shows

Source copy kept by wheresthe.ai for the story "Ransomware affiliates ran intrusions through AI agents, an exposed server shows" (<https://wheresthe.ai/d/intl-ransomware-affiliates-ran-intrusions-through-ai-agents-2026-10/>).

Source: <https://thehackernews.com/2026/10/exposed-infrastructure-reveals-how.html>

Text saved 24 Oct 2026 (the page as published).

SHA-256 of the saved text: c784c2aaa71aa4e8428f6615a45649334c3c0631e96d4fd946d1560ca9701d8f

Exposed Infrastructure Reveals How Ransomware Affiliates Used AI Agents to Execute Intrusions

Ravie Lakshmanan Oct 23, 2026 Ransomware / Cybercrime

Cybersecurity researchers have discovered an artificial intelligence (AI)-driven intrusion ecosystem run by affiliates of the INC and The Gentlemen ransomware groups.

Central to the operation is an "Autonomous Red Team Kill Chain Framework" referred to as Phantom. The threat actor is assessed to be Russian-speaking and functions both as an intrusion and data-extortion operator, as well as an infrastructure provider supporting different partners across the ransomware ecosystem.

"The operators exploited edge and remote-access systems at scale, including FortiClient EMS, SimpleHelp, cPanel, and VPN gateways," SOCRadar said. "They used AI-managed knowledge bases to track victims, stolen credentials, and campaign progress."

A notable aspect that defines the campaign is its "AI-executed, human-approved operating structure" in which a human operator provides a victim-specific prompt and reviews every step, while the intrusions themselves are executed through OpenCode with OpenRouter-hosted commercial models.

An examination of the exposed directory at 45.227.254[.]43 has revealed the activity logged 8,257 messages across 133 sessions. Should the models refuse to act on the prompts, the operator has been found to rewrite requests as an authorized red team assessment instead of simply switching models.

Initial access is achieved by exploiting public-facing edge devices at scale, following which the attackers move laterally, establish persistence, exfiltrate data, and escalate toward ransomware deployment. Operator records indicate that the attacks involve the exploitation of the following flaws -

CVE-2024-57727 (SimpleHelp path traversal) (Actively exploited)

CVE-2026-35616 (Fortinet FortiClient authentication bypass) (Actively exploited)

CVE-2026-41940 (cPanel/WHM authentication bypass) (Actively exploited)

CVE-2026-15409 (SonicWall SMA 1000 SSRF) (Staged for targeting, no confirmed compromise)

CVE-2026-63030 (WordPress wp2shell) (Developed)

CVE-2026-42208 (LiteLLM SQL injection) (Developed or referenced)

CVE-2026-9082 (Drupal SQL injection) (Developed or referenced)

"Once inside, the operators followed a recurring lifecycle: enumerate the domain, escalate to domain admin where possible, harvest credentials with stealer tooling, prioritize finance, HR, legal and executive files, and collect them over reverse-SOCKS tunnels," SOCRadar said.

"Credential reuse was the main force multiplier; in one case, a single reused password unlocked the VPN, servers, RMM, backup and security management platforms at once. Security products were impaired with a BYOVD killer and loader-level evasion, although the operators' own logs record mixed results against enterprise EDRs."

Supporting the threat actor in their goals is a diverse toolkit comprising bespoke frameworks, modified open-source tools, and compiled binaries -

Phantom, an AI-powered offensive security and penetration testing framework that's deployed in live operations via OpenCode and employs an agent that communicates with an external server to retrieve tasks, execute commands, and return outputs. It supports 18 skill modules to facilitate reconnaissance, binary reverse engineering, exploit research, steganography, telemetry assessment, and EDR fingerprinting, among others.

WormGPT-MCP, a set of 32 Model Context Protocol (MCP) servers connecting the AI system to 32 different modules that span functions ranging from reconnaissance to extortion.

Ghost C2, a collection of three generations of polymorphic implants that communicate with a remote server over Cloudflare, Telegram, or DNS, and execute tasking and maintain persistent beacons on compromised endpoints.

RevSocks, a reverse SOCKS5 framework used to pivot inside victim networks and stage data for exfiltration.

Other tools include a Cloudflare-fronted Mythic agent dubbed QuickLine, a Quick Heal/Seqrite EDR killer that abuses the Zemana driver ("zam64.sys") as part of a BYOVD attack, and exfiltration programs such as rks3 and BackupAgent.

Analysis of dozens of AI-authored extortion records shows that the threat actors adopt a data-centric approach. It involves examining the stolen financials to determine the amount a victim can pay without having to involve a bank or an insurer, and then sending masked proof samples and ramping up pressure through scripted vishing calls and threatening to notify customers, partners, and regulators.

The stolen data is eventually published on the data leak site if the victim fails to pay within a 24-72 hour deadline. That said, the final outcomes of these ransom demands are unknown as no cryptocurrency wallet or payment record has been recovered.

A total of 62 organizations are estimated to have been targeted by the operation, out of which four of them ended up being listed on the data leak sites for INC and The Gentlemen ransomware. The U.S. accounts for more than 50% of the targets. The attacks also primarily targeted industrial, manufacturing, and energy sectors, followed by healthcare and pharmaceutical organizations.

"Across this exposed infrastructure, the actors engaged in widespread exploitation of exposed internet services, specifically targeting FortiClient EMS gateways, Ivanti/Citrix/SonicWall/OpenVPN gateways, SimpleHelp remote management servers, and cPanel hosting controls," SOCRadar said.

"Concurrently, they orchestrated operations via the Phantom framework, established RevSocks reverse-SOCKS proxies for network pivoting, deployed Ghost C2 variants, and utilized AI-managed knowledge bases to meticulously log victim telemetry, compromised credentials, and active campaign statuses."

Found this article interesting? Follow us on Google News , Twitter and LinkedIn to read more exclusive content we post.

SHARE

Tweet

Share

Share

Share

SHARE

artificial intelligence , Cybercrime , data breach , Malware , network security , ransomware , Vulnerability

✂ Top Stories This Week

Researchers Demonstrate OpenBSD Root Exploit Chaining 37-Year-Old Shared-Memory Bug

Critical PwnEscape Linux Kernel Vulnerability Lets Guest VMs Read Host Memory

ThreatsDay: Adobe RCE, DNS Wallet Theft, NVIDIA Container Escape, Self-Healing Backdoor + 23 More

Citizen Lab Traces Zero-Click Candiru Spyware Infections at Al Jazeera to Targeted Online Ads

NGINX Fixes 5 Flaws, Including Possible RCE and a TLS Bug That May Skip Client-Cert Checks

Next.js Ships Out-of-Band Patch for Two Critical RCE Flaws in Caching and SVG Processing

Palo Alto Networks Patches PAN-OS Authenticated RCE Flaw Amid Active Exploitation

GhostAction Compromises About 1,900 GitHub Accounts, Targets Over 30,000 Repositories

Cisco Warns of Active Attacks Exploiting New Critical Secure Email Gateway Vulnerability

Langflow's Critical Authenticated RCE Flaw Targeted in 1,423 Honeypot Exploit Attempts

Hackers Backdoor Nextcloud Servers Through Critical Flaw in Collabora's Built-in CODE App

U.S. Shuts Down Over 40 Laptop Farms, Charges 14 in North Korean IT Worker Fraud

curl Patches HTTP/2 Upload Flaw After PoC Demonstrates Client-Side Code Execution

Researchers Used a Public OpenAI Model to Build a \$250,000 Chrome Full-Chain Exploit in Six Days

AI-Found Linux Kernel Flaw Gives Local Users Root, and Can Break a Guest Out to the Host

Microsoft Patches Record 1,413 Flaws, Including Exploited Windows Shell Bug

OpenAI Employee Arrested for Allegedly Planting Hidden Recorder in Executive Conference Room

Anthropic Warns Small AI Models Could Soon Form Self-Propagating IoT Botnets

Docker Fixes Critical Two-Year-Old Hub Flaw That Let One User Push Images as Another

SAP Discloses Critical CVSS 10 Flaw in Cloud SDK for Python Allowing OAuth2 Credential Theft

LG Smart TVs Vulnerable to Pre-Auth Root Takeover From Local Network

Attackers Heat French Supermarket Fridges Up to 50°C by Hijacking Danfoss Controllers

GitLab Patches CVSS 10.0 Flaw That Could Let Unauthenticated Users Run Server Code

⚡ Weekly Recap: Helpdesk 0-Days, Rogue Certs, ShinyHunters Arrest, AI-Fooling Malware and More

Researchers Publish Working Exploit for Pre-Auth AnyDesk Linux Flaw That Gives Root Access

★ Featured Resources

Download the 2026 Agent Security Report for a Five-Step Risk Reduction Framework

Get the Five-Stage Identity Maturity Roadmap for Governing AI Agents

Benchmark Your AI Risk Priorities Against 1,600 CISOs—Download the 2026 Report

No Exploit Yet? Preview the Summit Demo on Validating CVE Exposure