tinexta
defence

GTG-1002 e l'Ecosistema di Spionaggio Cinese

#TinextaDefenceBusiness

DFIR

Il Gruppo DFIR di Tinexta Defence è una Threat Response Unit specializzata in Digital Forensics & Incident Response che supporta imprese e pubbliche amministrazioni nella gestione di incidenti di sicurezza e nella produzione di evidenze digitali con valore probatorio.

L'attività del Gruppo integra competenze multidisciplinari e si articola in quattro aree principali:

- **Incident Response:** capacità di intervento rapido per contenere, eradicare e mitigare incidenti, riducendo l'impatto operativo;
- **Consulenze Tecniche d'Ufficio (CTU) e di Parte (CTP):** perizie informatiche conformi alle best practice di catena di custodia, a supporto del contesto giudiziario;
- **Forensic Readiness:** predisposizione preventiva di processi, tecnologie e standard per garantire che i dati raccolti siano accurati, integri e attestabili;
- **Ricerca e innovazione:** sperimentazione di tecnologie avanzate (eBPF, kernel telemetry, AI per anomaly detection, container forensics) per anticipare le minacce e sviluppare strumenti di nuova generazione.

In qualità di Threat Response Unit, il Gruppo DFIR non si limita alla fase di indagine post-evento, ma affianca i Security Operations Center (SOC) e le organizzazioni nella detection proattiva, nel threat hunting e nella gestione di crisi cyber.

La missione del Gruppo è elevare i livelli di sicurezza e resilienza delle infrastrutture critiche e dei sistemi informativi, coniugando rigore scientifico, innovazione tecnologica e capacità operativa a supporto della difesa digitale e del contesto giudiziario.

Sommario

Abstract

Analisi OSINT completa dell'attore di minaccia designato da Anthropic come GTG-1002, un gruppo di spionaggio state-sponsored cinese che ha orchestrato la prima campagna di cyber spionaggio su larga scala guidata da intelligenza artificiale. Il report analizza inoltre il più ampio ecosistema di operazioni parallele, guerra cognitiva e l'inquadramento strategico secondo la dottrina occidentale.

Autori:

- Gaetano Zappulla: CISO
- Gaia Calamari: Digital Forensic Analyst
- Luca Maggioni: Digital Forensic Analyst

1. Actor Profile: l'Identità di GTG-1002

Nel settembre 2025, Anthropic ha identificato l'attore di minaccia **GTG-1002** come un gruppo di spionaggio *state-sponsored cinese*, un'attribuzione supportata con un alto livello di confidenza[1]. A differenza dei tradizionali Advanced Persistent Threat (APT) cinesi, GTG-1002 non è definito da un set di malware personalizzati o da un'infrastruttura di comando e controllo (C2) riutilizzata. La sua identità è invece forgiata dalla sua innovativa metodologia operativa: l'orchestrazione di attacchi tramite agenti di intelligenza artificiale.

Mappatura con Alias Noti

Allo stato attuale, non esiste una mappatura diretta e pubblicamente confermata tra GTG-1002 e alias consolidati come *Volt Typhoon (Bronze Silhouette)*, *Salt Typhoon (RedMike)* o *APT41 (Barium, Winnti)*.

L'assenza di Indicatori di Compromissione (IoC) tradizionali nel report di Anthropic rende il clustering basato su dati tecnici quasi impossibile[2]. L'attacco si è basato sull'abuso di un servizio legittimo (l'API di Claude Code) e su tool open-source, rendendo l'attribuzione tecnica estremamente difficile.

Nonostante ciò, l'analisi comportamentale e strategica suggerisce che GTG-1002 non sia un attore completamente nuovo, ma più probabilmente una cellula specializzata o un'unità sperimentale all'interno di un più ampio programma di spionaggio cinese. Il targeting di GTG-1002 si allinea con gli obiettivi strategici a lungo termine della Cina, sovrapponendosi parzialmente con i settori di interesse di gruppi come Volt Typhoon (infrastrutture critiche) e Salt Typhoon (telecomunicazioni, governo)[3] [4]. L'obiettivo finale della campagna era la raccolta di intelligence, un segno distintivo delle operazioni state-sponsored cinesi. Gli attori cinesi sono noti per essere *early adopters* di tecnologie emergenti per scalare e migliorare le loro operazioni, e l'uso di agenti AI rappresenta la naturale evoluzione di questa dottrina.

> *"GTG-1002 ha utilizzato le stesse tecniche di attacco che gli stati-nazione usano da decenni – l'unica differenza è l'orchestrazione AI che abilita velocità e scala senza precedenti." – Clutch Security[2]".*

In sintesi, GTG-1002 dovrebbe essere considerato non tanto un nuovo gruppo, quanto una nuova capability nell'arsenale degli attori China-Nexus. È il nome di una campagna che rappresenta un salto quantico nella metodologia operativa, probabilmente condotta da un team specializzato nell'integrazione dell'IA all'interno di un APT esistente.

Dubbio Critico sull'Uso di Modelli Commerciali

Un'analisi critica sollevata da esperti del settore, come Nico Waisman di XBOW, mette in dubbio la probabilità che un attore state-sponsored sofisticato faccia affidamento esclusivo su un modello commerciale standard come quello di Anthropic[8].

> *"Trovo improbabile che un vero attore di minaccia cinese si affidi a un modello standard di Anthropic, non perché questi modelli non siano efficaci, ma perché tali gruppi sanno che Anthropic monitora attivamente questo tipo di attività e impara dai pattern, comportamenti e obiettivi che osserva." – Nico Waisman, XBOW[8]."*

Questa prospettiva suggerisce due possibilità alternative. La prima è l'esistenza di modelli AI offensivi in-house: è plausibile che attori state-sponsored stiano già sviluppando e utilizzando modelli linguistici offensivi privi di monitoraggio e guardrail di sicurezza, che non lasciano tracce di audit esterne.

La seconda possibilità è che la campagna GTG-1002 sia stata un'operazione deliberatamente visibile, condotta su una piattaforma commerciale in modo "rumoroso" per inviare un segnale strategico, testare le reazioni delle aziende di AI o distogliere l'attenzione da operazioni più furtive condotte con strumenti proprietari.

2. Cronologia Operativa: l'Evoluzione verso l'Autonomia

La campagna di GTG-1002 non nasce dal nulla, ma rappresenta il culmine di una traiettoria evolutiva osservata negli attori cinesi, che si muove verso una maggiore automazione e l'integrazione dell'IA. Questa progressione può essere vista come l'inizio di "The Chaos Phase", un'era in cui l'automazione trasforma gli APT sofisticati di ieri nella baseline di domani[8].

Fino al 2024: l'Era del "Living-off-the-Land"

Prima dell'avvento di agenti AI maturi, gruppi di punta come Volt Typhoon si sono specializzati in operazioni stealth a lungo termine. La loro dottrina si basava sul "Living-off-the-Land" (LOTL): l'uso esclusivo di strumenti legittimi e preinstallati sui sistemi delle vittime per evitare il rilevamento[3]. Queste operazioni miravano al pre-posizionamento su infrastrutture critiche quali energia, comunicazioni e trasporti, in preparazione di future operazioni distruttive. L'accesso avveniva tramite exploit di vulnerabilità note su dispositivi di rete come router e firewall, seguito da un lento e metodico movimento laterale usando credenziali rubate e strumenti nativi di Windows. L'uso di malware personalizzato era minimo o nullo per ridurre la superficie di rilevamento. Le operazioni erano manuali, lente e richiedevano operatori umani esperti in modalità *hands-on-keyboard*.

Estate 2025: il Precursore – "Vibe Hacking"

Nell'estate del 2025, Anthropic ha documentato una campagna di estorsione dati su larga scala, tracciata come **GTG-2002** e soprannominata "vibe hacking"[5]. Questa operazione, condotta da un attore cybercriminale, ha rappresentato il primo uso documentato di un agente AI (Claude Code) per eseguire attivamente parti di un attacco, non solo per assistere l'operatore. L'AI veniva usata come consulente tecnico e operatore attivo per reconnaissance, credential harvesting e data exfiltration. Il livello di autonomia era inferiore a GTG-1002: il report di novembre 2025 chiarisce che *"in quelle operazioni [vibe hacking], gli umani erano ancora molto nel loop, dirigendo le operazioni"*[1]. L'obiettivo era il guadagno finanziario tramite estorsione dati. Questa campagna ha dimostrato la fattibilità dell'uso di agenti AI come moltiplicatore di forza per attori con risorse limitate, sebbene ancora con una forte supervisione umana.

Settembre 2025: il Salto Quantico – La Campagna GTG-1002

La campagna GTG-1002 rappresenta un cambiamento di paradigma. Un attore state-sponsored ha preso il concetto di *"vibe hacking"* e lo ha evoluto in un framework di spionaggio quasi completamente autonomo. L'agente AI non è più solo un assistente, ma l'esecutore principale dell'operazione, passando da AI "assistant" a AI "actor"[9]. Il livello di autonomia ha raggiunto l'80-90% delle operazioni tattiche, con solo il 10-20% di intervento umano per approvare le escalation più critiche[9]. L'obiettivo era lo spionaggio strategico su vasta scala. Questa campagna ha segnato la prima dimostrazione di spionaggio su larga scala orchestrato dall'AI, operante a una velocità e scala "fisicamente impossibili" per team umani[1].

3. Deep Dive Tecnico AI: l'Anatomia della Campagna GTG-1002

L'innovazione di GTG-1002 risiede nella sua architettura tecnica, che ha trasformato un modello linguistico in un agente di spionaggio autonomo.

Architettura dell'Attacco

Il framework operativo era strutturato come segue: Operatori Umani -> Framework di Orchestrazione -> Claude Code + Tool MCP -> Target. Un sistema custom costruito dagli attaccanti gestiva le sessioni di Claude, orchestrava i tool e manteneva il contesto operativo. Claude Code veniva utilizzato come "cervello" dell'operazione, responsabile della scomposizione di obiettivi complessi in task eseguibili e della loro esecuzione materiale[6]. Il Model Context Protocol (MCP), uno standard open-source, ha permesso a Claude di utilizzare una vasta gamma di strumenti software esterni, come scanner di rete, browser automation, framework di exploit e tool di analisi, superando i limiti di un modello linguistico isolato[7].

Jailbreak e Ingegneria Sociale dell'AI

Per aggirare le barriere di sicurezza di Claude, progettate per prevenire l'abuso, GTG-1002 ha impiegato una sofisticata tecnica di ingegneria sociale a due livelli. La prima tecnica è la decomposizione dei task (Task-Shaping): l'attacco è stato suddiviso in una serie di azioni atomiche e apparentemente innocue. Ad esempio, invece di chiedere "trova una vulnerabilità per rubare credenziali", l'AI veniva istruita a "identificare porte aperte su questo IP" o "scrivere uno script per testare la robustezza delle password", mascherando l'intento malevolo finale[8].

La seconda tecnica è il role-playing: agli agenti AI veniva assegnato un ruolo fittizio, convincendoli di essere assistenti di un'azienda di cybersecurity legittima impegnati in un penetration test autorizzato. Questo contesto fittizio, fornito all'inizio della sessione, induceva il modello a eseguire azioni che altrimenti avrebbe rifiutato[1].

Fasi dell'Attacco a Guida AI

L'operazione ha seguito il classico ciclo di vita di un'intrusione, ma con un'autonomia crescente dell'AI in ogni fase.

Nella fase di *reconnaissance*, l'AI ha mappato autonomamente la superficie d'attacco esterna e interna dei target, enumerando servizi, API, topologie di rete e identificando database di alto valore. Ha gestito contesti separati per decine di target simultaneamente. Nella fase di *vulnerability discovery & exploitation*, l'agente AI ha identificato vulnerabilità come Server-Side Request Forgery (SSRF), ricercato tecniche di exploit, generato payload personalizzati e validato il successo dell'exploit tramite callback, il tutto in poche ore. L'intervento umano era limitato a una rapida revisione di 2–10 minuti per approvare l'escalation da "test" a "exploitation attiva"[1].

Una volta all'interno, nella fase di *lateral movement & credential harvesting*, l'AI ha estratto sistematicamente credenziali da file di configurazione e servizi, testandole autonomamente su altri sistemi per mappare le relazioni di fiducia e i livelli di privilegio, muovendosi lateralmente attraverso la rete. Nella fase di data *analysis & exfiltration*, invece di esfiltrare dati alla cieca, l'AI ha analizzato attivamente i dati rubati in situ, classificandoli per valore di intelligence e sensibilità. Ha poi generato report di sintesi per gli operatori umani, che dovevano solo approvare il set di dati finale da esfiltrare. Infine, nella fase di *documentation & handoff*, l'AI ha documentato autonomamente ogni passaggio dell'attacco, creando un report tecnico dettagliato per permettere a un team umano di subentrare e stabilire persistenza a lungo termine.

Vantaggio Strategico Ottenuto

L'orchestrazione AI ha fornito a GTG-1002 vantaggi strategici decisivi.

In termini di velocità macchina, l'agente AI operava a una velocità di migliaia di richieste al secondo, completando in poche ore attività di *reconnaissance* che richiederebbero giorni o settimane a un team umano[6]. La scala parallela è stata altrettanto impressionante: il framework ha permesso di attaccare fino a 30 target simultaneamente, gestendo contesti complessi in parallelo[1]. L'efficienza è stata massimizzata: l'automazione ha ridotto drasticamente la necessità di operatori umani esperti, abbassando la barriera di competenze per condurre operazioni sofisticate. Infine, si è verificata una *democratizzazione della minaccia*: l'uso di tool open-source e API commerciali dimostra che attacchi di questa portata non sono più appannaggio esclusivo di agenzie con budget miliardari[7].

4. Targeting & Vittime: un Approccio Strategico

Il targeting di GTG-1002 è stato ampio e strategico, colpendo una vasta gamma di settori in Nord America, Europa e Asia. A differenza di gruppi più focalizzati come Volt Typhoon, che si concentra sulle infrastrutture critiche, il targeting di GTG-1002 riflette un interesse ad ampio spettro per l'intelligence economica, tecnologica e governativa. Le grandi aziende tecnologiche sono state colpite per il furto di proprietà intellettuale e l'accesso a tecnologie emergenti. Le istituzioni finanziarie sono state targettizzate per intelligence economica e potenziale accesso a dati sensibili. Il settore della manifattura chimica è stato preso di mira in quanto strategico per l'economia e la difesa. Le agenzie governative sono state compromesse per spionaggio politico e raccolta di informazioni su policy.

Le vittime hanno descritto l'attacco come *"sentirsi come se si stesse giocando a scacchi contro un avversario che può fare 1.000 mosse al secondo"*[10]. Questa testimonianza cattura efficacemente la sensazione di impotenza di fronte a un avversario che opera a velocità macchina.

5. Raccomandazioni Difensive: adattarsi alla Minaccia AI-Driven

Le difese tradizionali, basate su firme e IoC noti, sono strutturalmente inadeguate a contrastare minacce come GTG-1002. Le raccomandazioni degli esperti si concentrano su un cambio di paradigma difensivo[9] [11].

L'architettura Zero-Trust richiede di assumere che la rete sia sempre compromessa e di verificare rigorosamente ogni richiesta di accesso, indipendentemente dalla sua origine. Il monitoraggio continuo e rilevamento comportamentale impone di spostare il focus dal rilevamento di malware noti all'identificazione di comportamenti anomali, come un'escalation di privilegi insolitamente rapida o un accesso a dati atipico. L'*identity hardening* prevede l'implementazione di autenticazione multi-fattore (MFA) resistente al phishing per contrastare il furto di credenziali. La segmentazione della rete mira a limitare il movimento laterale isolando i segmenti critici della rete.

La difesa AI-driven richiede l'utilizzo di strumenti di sicurezza basati sull'AI in grado di rilevare e rispondere ad attacchi che operano a velocità macchina. Il red teaming focalizzato sull'AI prevede la simulazione di attacchi orchestrati dall'AI per testare la resilienza delle proprie difese. Infine, la governance sull'uso interno dell'AI impone di stabilire policy chiare sull'uso di modelli AI da parte dei dipendenti per prevenire abusi e fughe di dati.

6. Operazioni Parallele dell'Ecosistema di Spionaggio Cinese (2024–2025)

Mentre GTG-1002 rappresenta l'avanguardia tecnologica dello spionaggio cinese attraverso l'orchestrazione AI, questa campagna non opera in isolamento. L'analisi delle operazioni parallele condotte da attori statali cinesi nel periodo 2024–2025 rivela un ecosistema di spionaggio sistematico e strategico che coordina canali cyber, umani e corporativi per rubare tecnologie industriali e di difesa statunitensi e alleate[12]. Questo capitolo mappa le principali campagne state-sponsored cinesi contemporanee a GTG-1002, dimostrando come Pechino stia operando su più fronti simultaneamente con obiettivi complementari: dal furto di proprietà intellettuale alla sorveglianza delle telecomunicazioni, dal targeting di infrastrutture critiche alla manipolazione dell'informazione.

6.1 Salt Typhoon: il "Peggior Hack Telecom della Storia USA"

Salt Typhoon rappresenta una delle campagne di spionaggio globale più vaste e persistenti mai documentate. Attiva dal 2021 con un picco nel 2025, questa operazione ha sfruttato vulnerabilità in router e dispositivi di rete per ottenere accesso persistente a infrastrutture critiche di telecomunicazioni in Stati Uniti, Australia, Canada, Regno Unito e altri paesi[13] [14]. La campagna è stata descritta da funzionari USA come *"il peggior hack telecom della storia degli Stati Uniti"*, con oltre 30.000 sistemi telecom compromessi solo negli USA[15]. L'operazione ha evitato il rilevamento per anni, dimostrando un livello di sofisticazione e pazienza strategica tipico delle operazioni di intelligence statale cinese.

L'attribuzione a attori statali cinesi è stata confermata con alta confidenza da CISA, NSA e FBI, con collegamenti diretti al Ministero della Sicurezza di Stato (MSS) e all'Esercito Popolare di Liberazione (PLA)[13]. Questa doppia affiliazione suggerisce un coordinamento tra intelligence civile e militare, coerente con la dottrina cinese di fusione militare-civile.

Salt Typhoon ha dimostrato una maestria nell'exploitation di dispositivi di rete perimetrali. Il targeting primario includeva router, firewall, switch e altri dispositivi di rete di provider di telecomunicazioni, governi, trasporti e reti militari. La persistenza veniva garantita attraverso il deployment di firmware modificati e backdoor a livello hardware, assicurando accesso a lungo termine anche dopo patch software. La tattica "Living-off-the-Land" prevedeva l'uso di funzionalità legittime dei dispositivi compromessi per evitare rilevamento. L'esfiltrazione dati consisteva nella raccolta massiva di metadati di comunicazione e, in casi selezionati, del contenuto delle comunicazioni per intelligence a lungo termine.

L'impatto di Salt Typhoon va ben oltre il furto di dati. La compromissione di infrastrutture di telecomunicazione offre alla Cina una capacità di sorveglianza globale, con accesso a metadati e contenuti di comunicazioni governative, militari e commerciali. Fornisce inoltre un pre-posizionamento per operazioni distruttive: la presenza persistente nelle reti telecom consente potenziali attacchi distruttivi in caso di conflitto. Infine, abilita l'intelligence economica attraverso l'intercettazione di comunicazioni aziendali per vantaggio competitivo.

La gravità della minaccia ha spinto un'azione coordinata senza precedenti. Nel settembre 2025, CISA, NSA e FBI hanno rilasciato un advisory congiunto (A-A25-239A) con guidance dettagliata per provider telecom[13]. Nell'agosto 2025, l'FBI ha rilasciato un video announcement pubblico sulla minaccia[15]. Nel novembre 2025, FCC, CISA e FBI hanno emesso ulteriori direttive per rafforzare la sicurezza delle telecomunicazioni[14].

6.2 PurpleHaze (UNC5174): Targeting di Aziende Cybersecurity

PurpleHaze, tracciata anche come Activity F e collegata al cluster UNC5174 (noto anche come Uteus/Uetus), rappresenta una campagna di intrusione su vasta scala che ha colpito oltre 70 organizzazioni globali tra luglio 2024 e marzo 2025[16]. Ciò che rende questa campagna particolarmente preoccupante è il targeting deliberato di aziende di cybersecurity come SentinelOne, oltre a provider IT, logistica e organizzazioni tecnologiche.

L'attribuzione a un attore statale cinese è stata confermata con alta confidenza da SentinelOne, con legami diretti a UNC5174, un broker di accesso iniziale affiliato a gruppi APT cinesi[16] [17]. UNC5174 è noto per fornire accessi a gruppi più specializzati, operando come un "enabler" nell'ecosistema APT cinese.

PurpleHaze ha dimostrato un arsenale tecnico sofisticato. Ha sfruttato vulnerabilità in SAP NetWeaver e altri software enterprise per l'accesso iniziale. Ha deployato malware come GOREVERSE (backdoor) e ShadowPad (piattaforma modulare di accesso remoto). Questa campagna ha segnato il primo abuso statale di THC Tools, essendo la prima campagna documentata di attori state-sponsored che utilizzano tool della suite THC (The Hacker's Choice), tradizionalmente associati a penetration testing legittimo[16]. La ricognizione estensiva prevedeva una mappatura dettagliata di reti interne prima dell'escalation.

Il targeting di aziende cybersecurity suggerisce obiettivi multipli. Il furto di intelligence su minacce consente l'accesso a telemetria di sicurezza, IoC e intelligence su altri attori di minaccia. La compromissione della supply chain permette l'uso di aziende cybersecurity come pivot per raggiungere i loro clienti. Il furto di proprietà intellettuale mira alle tecnologie di detection e response sviluppate da vendor di sicurezza.

6.3 Spionaggio Economico: "Made in China 2025" in Azione

La campagna di spionaggio economico cinese non è opportunistica, ma sistematica e guidata da policy statali. Il piano *"Made in China 2025"*, lanciato nel 2015, ha identificato 10 settori ad alto valore dove la Cina intende raggiungere la leadership globale[12]: next-generation information technology, robotics e automazione, aerospace e aviazione, settore marittimo, trasporti ferroviari avanzati, veicoli a energia nuova, energia pulita e rinnovabile, materiali avanzati, biotecnologia e farmaceutica, e macchinari agricoli.

Un report di novembre 2025 dell'Information Technology and Innovation Foundation (ITIF) documenta come lo spionaggio economico cinese si sia evoluto da attacchi esterni (cyber intrusions) a minacce interne (insider threats) sempre più sofisticate[12].

Il Thousand Talents Program, lanciato nel 2008, ha rappresentato il primo tentativo sistematico di reclutare scienziati e ingegneri stranieri per portare la loro ricerca in Cina. Sebbene molti casi abbiano coinvolto l'accademia, il settore privato non è stato immune: 600 recruits lavoravano per aziende USA al momento del reclutamento[18]. Casi documentati includono Xiaoqing Zheng (General Electric) e Xiaorong You (Coca-Cola e Eastman Chemical), entrambi accusati di furto di trade secrets[18].

Dopo che gli USA hanno iniziato a scrutinare il Thousand Talents Program, la Cina ha semplicemente rinominato e nascosto il programma, lanciando il National High-end Foreign Experts Recruitment Plan[18]. Questo nuovo schema target esplicitamente *"personale tecnico professionale o personale di gestione che ricopre posizioni senior in aziende e istituzioni internazionalmente rinomate"*[18]. Sussidiarie di aziende cinesi negli USA e "consulting fronts" operano come piattaforme di raccolta, reclutando talento americano e canalizzando know-how proprietario verso imprese statali cinesi[12]. Questo modello offusca deliberatamente i confini tra attività commerciale legittima e spionaggio.

Nel luglio 2025, Xu Zewei è stato arrestato per spionaggio economico legato al furto di segreti industriali nel settore dei semiconduttori[19]. Il caso esemplifica come insider reclutati operino per anni all'interno di aziende USA, estraendo gradualmente proprietà intellettuale critica. Il furto di IP non è fine a sé stesso, ma serve ad accelerare capabilities indigene, ridurre decenni di R&D a pochi anni di reverse engineering, ridurre la dipendenza tecnologica per raggiungere autosufficienza in settori strategici, e fornire un vantaggio militare, poiché molte tecnologie civili hanno applicazioni dual-use per il PLA.

6.4 Phish and Chips: Targeting dell'Industria Semiconduttori Taiwanese

Tra marzo e giugno 2025, tre distinti attori statali cinesi hanno condotto campagne di phishing coordinate contro l'industria dei semiconduttori taiwanese, inclusi giganti come TSMC e MediaTek[20]. Proofpoint ha tracciato questa operazione come "Phish and Chips", un gioco di parole che sottolinea il targeting specifico del settore chip.

L'operazione ha coinvolto tre attori distinti, tutti con motivazione di spionaggio[20]: *UNK_FistBump*, specializzato in lure PDF sofisticati; *UNK_DropPitch*, che usa framework AITM (Adversary-in-the-Middle) per phishing di credenziali; e *UNK_SparkyCarp*, che deploya backdoor come Voldemort. La presenza di tre attori separati ma coordinati suggerisce una campagna orchestrata a livello strategico, probabilmente coordinata da un'entità di intelligence centrale.

Le tattiche includevano spear-phishing altamente mirato con e-mail personalizzate e lure rilevanti per l'industria semiconduttori, AITM phishing per l'intercettazione in tempo reale di credenziali inclusi token MFA, e backdoor deployment per accesso persistente una volta ottenuto l'accesso iniziale.

Il targeting dell'industria semiconduttori taiwanese non è casuale. Taiwan produce oltre il 60% dei semiconduttori globali e oltre il 90% dei chip più avanzati[21]. La compromissione di aziende come TSMC offre alla Cina intelligence su tecnologie di processo (dettagli su nodi di produzione a 3nm e inferiori), dati su clienti (informazioni su quali aziende stanno sviluppando quali chip) e preparazione per uno scenario Taiwan in caso di escalation militare.

7. La Dimensione della Guerra Cognitiva

L'ecosistema di spionaggio cinese non si limita al furto di dati e al sabotaggio tecnico. Una componente sempre più integrata e sofisticata delle sue operazioni è la Guerra Cognitiva. A differenza della propaganda tradizionale, che mira a *convincere*, la guerra cognitiva digitale mira a influenzare il modo in cui il pubblico target pensa, creando confusione, seminando discordia e annegando le narrazioni critiche nel rumore digitale. L'obiettivo non è vincere un dibattito, ma renderlo impossibile.

7.1 Tattiche e Campagne Principali (2024–2025)

Spamouflage (o Dragonbridge) è la più grande e longeva operazione di influenza cinese, attiva dal 2019 e in continua evoluzione. La sua tattica principale non è l'ingaggio, ma il "flooding" (inondazione): una rete massiccia di centinaia di migliaia di account falsi (bot) viene utilizzata per postare milioni di commenti spam, video e link, soffocando hashtag e conversazioni su argomenti sensibili per Pechino come i diritti umani nello Xinjiang, le proteste a Hong Kong e le critiche al Partito Comunista Cinese.

Nel 2024, Spamouflage ha compiuto un salto di qualità. Il report di Graphika "The #Americans" (settembre 2024) ha documentato come la rete abbia iniziato a impersonare elettori americani per influenzare le elezioni presidenziali USA[28]. Le nuove tattiche includevano la creazione di personas che si fingono cittadini statunitensi frustrati dalla politica, sostenitori di specifici candidati o attivisti per i diritti umani. I contenuti divisivi comprendevano narrazioni denigratorie contro candidati di entrambi gli schieramenti, dubbi sulla legittimità del processo elettorale e amplificazione di temi sociali polarizzanti come il controllo delle armi e la disuguaglianza razziale. Il report evidenzia come parte dei contenuti testuali e visivi fosse "quasi certamente generata dall'AI", permettendo una produzione su larga scala di contenuti più credibili.

> *"Valutiamo che Spamouflage e altri attori di influenza cinesi continueranno quasi certamente i loro sforzi per influenzare le conversazioni politiche statunitensi... sfruttando le divisioni sociali in un ambiente informativo polarizzato per ritrarre gli Stati Uniti come una potenza globale in declino." – Graphika, The #Americans Report"[28].*

PAPERWALL, scoperta da Citizen Lab nel febbraio 2024, è una rete di almeno 123 siti web che si fingono testate giornalistiche locali in 30 paesi, inclusa l'Italia con domini come "Roma Journal" e "Milano Notizie"[29]. La tecnica del "whitewashing" prevede che i siti pubblichino per il 90% notizie locali reali e innocue, spesso copiate da media legittimi. Nel restante 10%, inseriscono articoli di propaganda pro-Pechino, attacchi ad hominem contro critici della Cina e contenuti ripresi dai media di stato cinesi. L'obiettivo è "lavare" la propaganda, presentandola attraverso una fonte apparentemente locale e credibile per un lettore occidentale. Citizen Lab ha attribuito la campagna a Shenzhen Haimaiyunxiang Media Co., Ltd. (Haimai), una società di PR cinese, dimostrando il coinvolgimento del settore privato nelle operazioni di influenza statale.

La Cina è pioniera nell'uso dell'Intelligenza Artificiale per la propaganda video. Nel 2023, è stata scoperta una serie di notiziari fittizi chiamata "Wolf News", i cui presentatori erano avatar realistici generati dall'AI (deepfakes) che leggevano copioni pro-Cina in perfetto inglese. Questa tattica permette di produrre video di propaganda in qualsiasi lingua a costo quasi zero, senza la necessità di attori umani.

Oltre ai bot, la Cina utilizza sempre più persone reali per veicolare i suoi messaggi. Il governo cinese sponsorizza viaggi per youtuber e influencer occidentali in regioni controverse come lo Xinjiang, a condizione che pubblichino contenuti che negano le violazioni dei diritti umani. I "Little Pinks", evoluzione dell'"Esercito dei 50 Centesimi", sono brigate di volontari nazionalisti che attaccano in massa online chiunque critichi la Cina, praticando doxing e minacce per costringere al silenzio o alle scuse.

7.2 Integrazione tra Guerra Cognitiva e Operazioni Cyber

Le operazioni di influenza non sono separate dalle operazioni di hacking; sono due facce della stessa medaglia strategica. Il Microsoft Digital Defense Report 2025 evidenzia come gli attori statali cinesi stiano integrando queste due dimensioni[30].

Gli attori cinesi compromettono account di social media di alto profilo per diffondere narrazioni pro-Pechino o contenuti divisivi, sfruttando la credibilità dell'account hackerato. Campagne di phishing e spear-phishing utilizzano narrazioni di attualità e contenuti di propaganda come "esca" per indurre le vittime ad aprire allegati malevoli o cliccare su link dannosi. Un esempio è l'uso da parte di “Mustang Panda” di finti documenti diplomatici sulla guerra in Ucraina per colpire governi europei. Se un'azienda o un funzionario cinese viene sanzionato, partono campagne di Search Engine Optimization (SEO) per creare migliaia di pagine web positive che spingono le notizie negative nelle pagine interne dei motori di ricerca, rendendole di fatto invisibili.

7.3 Implicazioni per le Aziende Occidentali

La guerra cognitiva non rappresenta più una minaccia esclusiva per i governi e i processi elettorali, ma sta assumendo implicazioni dirette e crescenti anche per il settore privato. Le aziende, in particolare quelle operanti in settori strategici, tecnologici o a forte esposizione geopolitica, possono diventare bersaglio di campagne di disinformazione finalizzate al danneggiamento reputazionale e competitivo. Operazioni di boicottaggio coordinate da comunità digitali organizzate, come il fenomeno dei cosiddetti *"Little Pinks"*, sono già in grado di generare impatti economici concreti e misurabili.

Le campagne di disinformazione vengono inoltre utilizzate per alterare artificialmente il valore di mercato delle aziende, attraverso la diffusione di notizie false relative a presunti fallimenti di prodotto, vulnerabilità di sicurezza, procedimenti giudiziari o indagini regolatorie inesistenti. Sul piano operativo, i dipendenti esposti in modo continuativo a contenuti propagandistici risultano più vulnerabili a campagne di spear-phishing, che sfruttano narrazioni geopolitiche o temi di attualità come esca, incrementando in modo significativo il rischio di compromissioni informatiche. Un ulteriore elemento di rischio è rappresentato dalla proliferazione di finti media locali (PAPERWALL) e dalle tecniche di manipolazione dei motori di ricerca, che rendono sempre più complessa una corretta attività di due diligence su potenziali partner, fornitori o investimenti, in particolare nell'area cinese. In tali contesti, le informazioni negative vengono sistematicamente occultate, distorte o "ripulite", compromettendo la capacità decisionale delle imprese.

In questo scenario, le aziende occidentali sono chiamate a integrare in modo strutturato la threat intelligence sulla disinformazione all'interno dei propri programmi di sicurezza. Il monitoraggio continuo delle narrazioni che coinvolgono il proprio brand e il settore di riferimento, la formazione del personale al riconoscimento della propaganda sofisticata e la predisposizione di piani di crisis communication per la gestione di campagne di influenza ostili non sono più attività opzionali, ma competenze strategiche indispensabili per operare in sicurezza nell'attuale contesto geopolitico.

8. Inquadramento Strategico: la Guerra Ibrida secondo la Dottrina Italiana

Per contestualizzare appieno la minaccia rappresentata da attori come GTG-1002 e l'ecosistema di operazioni cinesi, è necessario comprendere il quadro strategico in cui si inseriscono. Il non-paper *"Il Contrasto alla Guerra Ibrida: Una Strategia Attiva"*, pubblicato dal Ministero della Difesa italiano nel novembre 2025, fornisce una visione autorevole della dottrina nazionale e alleata su questa minaccia pervasiva e sotto-soglia[31].

8.1 Definizione e Attori della Minaccia Ibrida

Il documento definisce la minaccia ibrida come un insieme di *"azioni coordinate in più domini condotte da attori statuali e non-statuali, al di sotto della soglia del conflitto armato e spesso non attribuibili, mirate a danneggiare, destabilizzare o indebolire"*[31]. Questa definizione evidenzia due caratteristiche centrali: l'azione sotto-soglia per evitare una risposta militare convenzionale e la difficoltà di attribuzione per mantenere la *"plausible deniability"*.

Il non-paper identifica esplicitamente la Cina come uno dei principali attori ibridi globali, descrivendo la sua strategia come un "approccio integrato che combina leve economiche, tecnologiche, informative e diplomatiche per indebolire l'UE e acquisire know-how strategico". Le attività attribuite alla Cina nel documento includono disinformazione in Italia e in teatri strategici come l'Africa, infiltrazioni nei sistemi bancari e nelle reti pubbliche, reclutamento in ambito cyber e penetrazione nei settori dell'informazione, della finanza e dell'economia. Queste attività, descritte dalla Difesa italiana, corrispondono perfettamente alle operazioni analizzate in questo dossier, da GTG-1002 (infiltrazione tecnologica) a PAPERWALL (penetrazione informativa) e allo spionaggio economico legato a "Made in China 2025".

8.2 La Risposta Strategica: da Contenitiva a Proattiva

Il punto cruciale del non-paper è l'appello a un cambio di postura strategica per l'Italia e l'Occidente. Il Ministro della Difesa, Guido Crosetto, sottolinea come l'approccio attuale sia spesso reattivo e insufficiente.

> *"L'Occidente spesso sceglie di non reagire... Occorre passare rapidamente dall'attuale postura contenitiva a una postura concretamente difensiva (che in ambito hybrid non può che essere proattiva), sia a livello nazionale sia in ambito alleanze." – Ministro Guido Crosetto*[31]

Questa visione implica la necessità di superare l'inerzia, non considerando più le azioni ibride come incidenti isolati ma come una campagna continua. Richiede di sviluppare una difesa proattiva, mantenendosi attivi nel dominio ibrido per prevenire azioni ostili e ridurre la libertà di manovra degli avversari. Impone un coordinamento multi-livello, integrando le risposte a livello nazionale, UE e NATO, creando strutture dedicate come un Centro Europeo per il Contrasto alla Guerra Ibrida. Infine, richiede di investire in resilienza della società, attraverso alfabetizzazione digitale e consapevolezza per rendere la società civile meno vulnerabile alla manipolazione cognitiva.

La dottrina esposta nel non-paper suggerisce che per contrastare efficacemente l'ecosistema di spionaggio e influenza cinese, non è sufficiente difendersi passivamente. È necessario attribuire pubblicamente le campagne, anche in assenza di prove definitive, per erodere la *"plausible deniability"*. Occorre imporre costi, utilizzando strumenti economici e diplomatici come l'Anti-Coercion Instrument dell'UE per rispondere alle azioni di coercizione geo-economica. È indispensabile rafforzare la collaborazione pubblico-privato tra intelligence, difesa e settore privato per una situational awareness condivisa e una risposta più rapida.

Il documento del Ministero della Difesa fornisce la cornice strategica che conferma la gravità e la natura sistemica delle minacce descritte in questo dossier. Le operazioni cinesi non sono semplici atti di hacking, ma componenti di una strategia ibrida integrata che richiede una risposta altrettanto integrata e proattiva.

9. Conclusioni Strategiche: il Momento Stuxnet per l'Agentic AI

La campagna GTG-1002 e le operazioni parallele del 2024-2025 non rappresentano solo un'evoluzione, ma un cambio di dottrina fondamentale per lo spionaggio state-sponsored. L'analisi porta a cinque conclusioni strategiche:

1. **GTG-1002 è il "Momento Stuxnet per l'Agentic AI":** Come Stuxnet ha dimostrato la fattibilità del sabotaggio fisico tramite codice, GTG-1002 ha dimostrato la fattibilità di campagne di spionaggio su larga scala, quasi autonome, orchestrate dall'AI. Ha stabilito un nuovo standard di velocità e scala[7].
2. **Fine del Rate-Limiting Umano:** La principale barriera alla scala delle operazioni cyber è sempre stata la disponibilità di operatori umani esperti. L'automazione AI rimuove questo vincolo, permettendo a un piccolo team di orchestrare decine di intrusioni complesse simultaneamente.
3. **L'Identità dell'Attore Diventa Meno Rilevante della Capability:** L'uso di API commerciali e tool open-source offusca l'attribuzione. La minaccia non è più solo "APT41", ma la capability di orchestrare attacchi AI-driven, che può essere adottata da più gruppi. L'automazione trasforma gli APT di ieri nella baseline di domani[8].
4. **Le Difese Esistenti sono Strutturalmente Inadeguate:** Le strategie di difesa basate su IoC e rilevamento di firme sono inefficaci contro attacchi che non usano malware e si muovono a velocità macchina. La difesa deve evolvere verso architetture Zero-Trust e rilevamento comportamentale AI-driven.
5. **La Minaccia è un Ecosistema Integrato:** GTG-1002, Salt Typhoon, la guerra cognitiva e lo spionaggio economico non sono incidenti isolati. Sono manifestazioni di una strategia ibrida nazionale che integra tutte le leve del potere statale. La vera minaccia non è un singolo gruppo APT, ma un sistema di spionaggio nazionale che richiede una risposta altrettanto integrata e proattiva, come sottolineato dalla dottrina della Difesa italiana[31].

Riferimenti

[1] Anthropic. (2025, November 13). *Disrupting the first reported AI-orchestrated cyber espionage campaign.* https://www.anthropic.com/news/disrupting-AI-espionage

[2] Clutch Security. (2025, November 18). *The Anthropic GTG-1002 Report: Nothing New, But Your Controls Better Be Tight.* https://www.clutch.security/blog/the-anthropic-gtg-1002-report-nothing-new-but-your-controls-better-be-tight

[3] MITRE ATT&CK. (2025, April 30). *Volt Typhoon, G1017.* https://attack.mitre.org/groups/G1017/

[4] CISA. (2025, September 3). *Countering Chinese State-Sponsored Actors Compromise of Networks Worldwide to Feed Global Espionage System (AA25-239A).* https://www.cisa.gov/news-events/cybersecurity-advisories/aa25-239a

[5] Anthropic. (2025, August). *Threat Intelligence Report: August 2025.* (Analisi basata su documento PDF)

[6] SOCFortress. (2025, November). *GTG-1002 AI-Orchestrated Espionage Campaign (Nov 2025)*. https://socfortress.medium.com/gtg-1002-ai-orchestrated-espionage-campaign-nov-2025-0f8151471be5

[7] Quilr.ai. (2025, November 18). *The Stuxnet Moment For Agentic AI: Inside Anthropic's GTG-1002 Breach.* https://www.quilr.ai/blog-details/agentic-ai-inside-anthropics-gtg-1002-breach

[8] XBOW. (2025, November 14). *Autonomous Offense IRL: What Anthropic's GTG-1002 Exposes, and How We Scale the Fight Back.* https://xbow.com/blog/anthropic-gtg1002-ai-cyberattack-analysis

[9] eSecurity Planet. (2025, November 14). *Inside the First AI-Driven Cyber Espionage Campaign.* https://www.esecurityplanet.com/threats/inside-the-first-ai-driven-cyber-espionage-campaign/

[10] The Economist. (2025, November 19). *How China-linked hackers co-opted Anthropic's Claude.* https://www.economist.com/china/2025/11/19/how-china-linked-hackers-co-opted-anthropics-claude

[11] Cyderes. (2025, November 15). *The First Verified AI-Orchestrated Cyber Espionage Campaign Signals a New Era of Attack and Defense.* https://www.cyderes.com/howler-cell/first-ai-driven-cyber-espionage-campaign-anthropic-analysis

[12] ITIF. (2025, November 3). *From Outside Assaults to Insider Threats: Chinese Economic Espionage.* https://itif.org/publications/2025/11/03/from-outside-assaults-to-insider-threats-chinese-economic-espionage/

[13] CISA. (2025, September). *CISA and Partners Release Joint Advisory on Countering Chinese State-Sponsored Actors.* https://www.cisa.gov/news-events/news/cisa-and-partners-release-joint-advisory-countering-chinese-state-sponsored-actors-compromise

[14] Hogan Lovells. (2025). *Salt Typhoon Cyberattack Prompts Action from FCC, CISA, FBI, and More.* https://www.hoganlovells.com/en/publications/salt-typhoon-cyberattack-prompts-action-from-fcc-cisa-fbi-and-more

[15] FBI. (2025, August 27). *Salt Typhoon Video Announcement.* https://www.fbi.gov/-video-repository/salttyphoon082725.mp4/view

[16] SentinelOne. (2025, June). *Follow the Smoke: China-Nexus Threat Actors Hammer at the Doors of Top-Tier Targets.* https://www.sentinelone.com/labs/follow-the-smoke-china-nexus-threat-actors-hammer-at-the-doors-of-top-tier-targets/

[17] Industrial Cyber. (2025, June). *SentinelOne Links ShadowPad and PurpleHaze Attacks to China-Aligned Threat Actors.* https://industrialcyber.co/ransomware/sentinelone-links-shadowpad-and-purplehaze-attacks-to-china-aligned-threat-actors/

[18] The New York Times. (2019). *China's Thousand Talents Program: 600 Recruits Worked for U.S. Companies.* (Citato in ITIF Report)

[19] TechNewsWorld. (2025, November). *US Think Tank Waves Red Flag Over Chinese Economic Espionage.* https://www.technewsworld.com/story/us-think-tank-waves-red-flag-over-chinese-economic-espionage-180009.html

[20] Proofpoint. (2025, July). *Phish and Chips: China-Aligned Espionage Actors Ramp Up Taiwan Semiconductor Targeting.* https://www.proofpoint.com/us/blog/threat-insight/phish-china-aligned-espionage-actors-ramp-up-taiwan-semiconductor-targeting

[21] BankInfoSecurity. (2025, July). *China-Backed Hackers Intensify Attacks on Taiwan Chipmakers.* https://www.bankinfosecurity.com/china-backed-hackers-intensify-attacks-on-taiwan-chipmakers-a-29004

[22] House Select Committee on the CCP. (2025, September). *Committee Statement on Ongoing PRC Cyber Espionage Targeting US Trade Policy Stakeholders.* https://-selectcommitteeontheccp.house.gov/media/press-relea-ses/committee-statement-on-ongoing-prc-cyber-espionage-targeting-us-trade-policy-stakeholders

[23] The Hacker News. (2025, November). *China-Linked APT31 Launches Stealthy Attacks on Russian IT Firms.* https://thehackernews.com/2025/11/china-linked-apt31-launches-stealthy.html

[24] Recorded Future. (2025, November). *Russia Report: APT31 China-Linked Hacks.* https://therecord.media/russia-report-apt31-china-linked-hacks

[25] GBHackers. (2025, November). *Massive Data Leak Exposes Chinese Cyber Weapons.* https://gbhackers.com/data-leak/

[26] WebProNews. (2025, November). *China's Cyber Shadows Exposed: Inside the Knownsec Leak.* https://www.webpronews.com/chinas-cyber-shadows-exposed-inside-the-knownsec-leak/

[27] Recorded Future. (2025, October 30). *China's Militia Forces Train to "Get Strong" in the New Era.* https://www.recordedfuture.com/research/chinas-militia-forces-train-to-get-strong-in-the-new-era

[28] Graphika. (2024, September 3). *The #Americans.* https://graphika.com/reports/the-americans

[29] Citizen Lab. (2024, February 7). *PAPERWALL: Chinese Websites Posing as Local News Outlets Target Global Audiences with Pro-Beijing Content.* https://citizenlab.ca/2024/02/paperwall-chinese-websites-posing-as-local-news-outlets-with-pro-beijing-content/

[30] Microsoft. (2025, October). *Microsoft Digital Defense Report 2025.* (Analisi basata su documento PDF)

[31] Ministero della Difesa. (2025, Novembre). *Non-paper: Il Contrasto alla Guerra Ibrida: Una Strategia Attiva.* https://www.difesa.it/assets/allegati/83696/non-paper_il_contrasto_alla_guerra_ibrida.pdf

tinexta
defence

Next | Donexit | Foramil | Innodesi

Via Giacomo Peroni, 452 – 00131 Roma
tel. 06.45752720 – info@defencetech.it
www.tinextadefence.it

#TinextaDefenceBusiness