tinexta
defence

Deception Technology e Moving Target Defense:
Strategia di Sicurezza Attiva per l'Impresa Moderna

#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

Executive Summary

*Nel 2025, un attaccante permane mediamente 241 giorni all'interno delle reti aziendali, con un impatto significativo su rischio operativo, costi e continuità del business. Questo report illustra come ridurre radicalmente tale permanenza – con un obiettivo inferiore alle 24 ore – attraverso un cambiamento di paradigma: dal modello di sicurezza passivo e reattivo a una **Sicurezza Attiva (Active Defense)** basata su **Deception Technologies** e **Moving Target Defense (MTD)**. La visione proposta integra tecnologia, processi e governance per trasformare la cybersecurity in una capacità proattiva di resilienza.*

In un panorama dove la prevenzione totale è un'illusione e gli attaccanti operano indisturbati all'interno delle reti per mesi, l'adozione di un approccio proattivo che caccia attivamente l'avversario rappresenta non solo un'opportunità tecnologica, ma una necessità strategica. Questa visione è validata da framework di settore come MITRE Engage e analisti di mercato come Gartner[9] [11].

Le piattaforme di Deception trasformano l'infrastruttura in un ambiente ostile per l'attaccante, nel quale ogni tentativo di esplorazione o abuso di risorse fittizie produce evidenze ad elevata affidabilità. Rispetto ai meccanismi di detection tradizionali, consentono una significativa riduzione degli eventi non qualificati e permettono di individuare le compromissioni nelle fasi iniziali del ciclo di attacco, riducendo sensibilmente i tempi di permanenza dell'avversario all'interno dei sistemi. Secondo il report IBM "Cost of a Data Breach 2025", la riduzione del Dwell Time rappresenta uno dei principali fattori di mitigazione dell'impatto economico degli incidenti: nel 2025 il costo medio globale di un data breach si attesta a 4,44 milioni di dollari, con un tempo medio di permanenza dell'attaccante pari a 241 giorni, il livello più basso registrato nell'ultimo decennio[1].

In sinergia, le tecnologie MTD e la loro evoluzione nell'Automated Moving Target Defense (AMTD) rendono l'infrastruttura un bersaglio mobile e imprevedibile, aumentando esponenzialmente i costi e la complessità per l'attaccante. Gartner definisce l'AMTD una tecnologia ad alto impatto strategico[11] e prevede che le tecnologie di sicurezza preemptive, incluse Deception e MTD, costituiranno oltre il 50% della spesa in IT security entro il 2030, rispetto a meno del 5% nel 2024[11].

Questo documento vuole fornire una analisi tecnica approfondita, arricchita da studi accademici *peer-reviewed*[14] [16], raccomandazioni governative[17], una roadmap strategica per l'integrazione di queste tecnologie, i KPI per misurarne il successo e un solido business case fondato su dati di mercato e standard di settore per massimizzare il ritorno sull'investimento (ROI) e costruire una difesa resiliente e adattiva.

Autori:

- Gaetano Zappulla: Chief Information Security Officer
- Michele Fredella: Cybersecurity Research Analyst
- Giorgia Lorusso: Sostituto Punto di Contatto NIS2

1. Il Contesto: Il Fallimento della Prevenzione Tradizionale

Il tradizionale modello di sicurezza informatica basato sulla difesa perimetrale risulta oggi strutturalmente inadeguato rispetto all'evoluzione degli ecosistemi digitali. Tale approccio presuppone l'esistenza di un confine chiaro tra interno ed esterno dell'organizzazione e si fonda su un principio di fiducia implicita verso ciò che risiede all'interno del perimetro. Questo presupposto non è più valido. La progressiva adozione di cloud computing, architetture ibride e multi-cloud, mobilità, lavoro remoto e servizi digitali distribuiti ha determinato una frammentazione del perimetro tradizionale, rendendo inefficaci le difese progettate per ambienti statici e centralizzati. Firewall, antivirus e sistemi di prevenzione perimetrale, pur rimanendo componenti necessarie, non sono più sufficienti a garantire un livello di sicurezza adeguato. Numerosi studi di settore evidenziano come gli attaccanti, una volta ottenuto un accesso iniziale – spesso tramite credenziali legittime o tecniche di *social engineering* – siano in grado di muoversi lateralmente all'interno delle infrastrutture per periodi prolungati, eludendo i controlli tradizionali e operando indisturbati sui flussi di traffico est-ovest.

In questo contesto, emerge la necessità di un cambio di paradigma: dall'obiettivo irrealistico di prevenire ogni possibile compromissione a un modello operativo basato sull'assunzione della compromissione (*assume breach*). La questione strategica non è più se un attaccante riuscirà a entrare, ma quando, e la priorità diventa la capacità di rilevare tempestivamente comportamenti anomali e attività malevole una volta che l'attaccante è già all'interno dell'ambiente digitale.

Questo approccio è coerente con le raccomandazioni dei principali framework di riferimento europei e internazionali, che sottolineano l'importanza di capacità avanzate di detection, visibilità interna e risposta rapida, come elementi fondamentali per ridurre l'impatto degli incidenti e garantire la resilienza operativa delle organizzazioni.

Il problema principale degli Intrusion Detection Systems (IDS) tradizionali è l'alto tasso di falsi allarmi, che genera un progressivo sovraccarico operativo nei team di sicurezza[16]. L'uso di honeypot e tecnologie di deception fornisce una soluzione efficace a questo problema, distinguendo chiaramente tra traffico legittimo e attività malevole, riducendo significativamente l'overhead operativo[16].

Categoria di Minaccia	**Debolezza della Difesa Tradizionale**	**Copertura Deception/MTD**
Attacchi Zero-Day	Le difese basate su firme (Antivirus, IPS) sono intrinsecamente inefficaci.	MTD mitiga l'impatto; *Deception* cattura *exploit attempts*.
Movimento Laterale	Scarsa visibilità del traffico Est-Ovest, che costituisce la maggior parte dell'attività in un attacco avanzato.	*Breadcrumbs e Decoy* intercettano *lateral movement* (MITRE ATT&CK TA0008).
Insider Threat	Incapacità di distinguere tra comportamento legittimo e malevolo di un utente con credenziali valide. Gli attacchi *insider* costano in media $4.99 milioni per incidente (IBM 2025), il 12% in più rispetto alla media globale.	*Honey credentials* rilevano abuso di privilegi.

2. L'Evoluzione dell'Inganno: dagli Honeypot alle Piattaforme Distribuite

2.1 Deception in Ambienti Cloud-Native

La progressiva adozione di architetture cloud-native, basate su virtualizzazione, containerizzazione e servizi gestiti, modifica in modo sostanziale anche le modalità con cui la Deception Technology deve essere progettata e implementata. L'inganno non può più essere concepito come un insieme di honeypot statici collocati in una rete tradizionale, ma deve diventare una componente nativa dell'infrastruttura software-defined, capace di adattarsi dinamicamente ai cambiamenti dell'ambiente. Negli ambienti *Infrastructure-as-a-Service (IaaS)*, la deception viene realizzata distribuendo macchine virtuali esca che replicano fedelmente le caratteristiche degli asset reali. Negli ambienti containerizzati e nelle piattaforme Kubernetes, la Deception assume la forma di workload esca, namespace trappola, segreti e token fasulli. Nei modelli *Software-as-a-Service e Identity-as-a-Service*, la deception opera sul piano identitario e applicativo attraverso account esca, token API fittizi e documenti trappola. Hyperscaler come AWS hanno implementato sistemi di deception su scala globale (MadPot) per proteggere la propria infrastruttura, rilevando e neutralizzando centinaia di minacce quotidianamente e condividendo intelligence con l'ecosistema di sicurezza[21].

2.2 Le Origini: Gli Honeypot

Il concetto di usare esche per ingannare gli attaccanti ha radici profonde nella storia della cybersecurity, ma la sua implementazione si è evoluta radicalmente nel corso dei decenni.

La storia degli honeypot inizia con il concetto di *"entrapment"* (intrappolamento), definito già nel 1976 dallo standard FIPS 39 come "il deliberato inserimento di apparenti falle in un sistema allo scopo di rilevare tentativi di penetrazione o confondere un intruso su quali falle sfruttare"[5]. Tuttavia, la prima applicazione pratica documentata risale al 1986, quando l'astronomo Clifford Stoll, al Lawrence Berkeley National Laboratory, scoprì un'anomalia contabile di 75 centesimi nel tempo di calcolo. Invece di chiudere la falla, Stoll creò un elaborato sistema di file e dati fittizi per tenere l'attaccante impegnato, riuscendo a tracciare per quasi un anno uno spy ring al soldo del KGB. La sua storia è raccontata nel celebre libro *"The Cuckoo's Egg"* (1989)[6], che di fatto rappresenta il primo trattato sulla cyber-spionaggio e l'uso di honeypot per la contro-intelligence.

Pochi anni dopo, nel gennaio 1991, Bill Cheswick presso gli AT&T Bell Labs costruì una delle prime *"prigioni"* digitali, un *"chroot jail"*, per osservare un hacker che tentava di rubare file di password, monitorandolo per diversi mesi[7]. Questi primi honeypot erano sistemi artigianali, manuali e ad alta intensità di lavoro, ma dimostrarono il valore strategico dell'inganno come strumento difensivo.

Il passo successivo in questa evoluzione fu la transizione da sforzi individuali a una ricerca comunitaria e organizzata. Questo momento è segnato dalla nascita, nel 1999, del Honeynet Project, un'organizzazione di ricerca non-profit internazionale fondata da Lance Spitzner. L'obiettivo del progetto era, ed è tuttora, quello di 'imparare gli strumenti, le tattiche e le motivazioni della comunità blackhat' e condividere queste conoscenze. Il Honeynet Project ha avuto un ruolo fondamentale nel formalizzare l'uso degli honeypot come strumenti di ricerca su larga scala, sviluppando strumenti open-source e promuovendo un approccio sistematico alla raccolta di intelligence sulle minacce.

2.3 Classificazione e Sviluppo

Con il tempo, gli honeypot si sono evoluti e sono stati classificati principalmente secondo due criteri: obiettivo e livello di interazione. La ricerca accademica ha ulteriormente raffinato questa tassonomia, introducendo un terzo criterio di specializzazione per affrontare minacce specifiche[18].

Criterio	Tipologia	Descrizione	Vantaggi	Svantaggi
Obiettivo	*Production Honeypot*	Rilevare intrusioni nella rete aziendale e deviare gli attacchi. Generano alert per il SOC.	Protezione diretta degli asset.	Focus su *detection*, non su ricerca.
	Research Honeypot	Studiare le TTP degli attaccanti e raccogliere campioni di *malware*. Usati da accademici e *threat intelligence analyst*.	*Intelligence* approfondita.	Non protezione diretta.
Interazione	*Low-Interaction*	Emulano servizi e protocolli. L'attaccante non interagisce con un sistema operativo reale.	Economici, facili da implementare, basso rischio.	Facilmente smascherabili da attaccanti sofisticati.
	High-Interaction	Forniscono un sistema operativo reale (spesso virtualizzato) con cui l'attaccante può interagire pienamente.	Difficili da rilevare, *intelligence* dettagliata.	Costi elevati, rischio di compromissione completa.
Specializzazione	*APT Honeypot*	Progettati per rilevare campagne *Advanced Persistent Threat*.	Rilevamento di minacce sofisticate.	Complessità di configurazione.
	Ransomware Honeypot	Rilevano attività di crittografia e tracciano pagamenti (*Bitcoin Honeytoken*).	*Early warning* su attacchi *ransomware*.	Richiede aggiornamenti costanti.
	Insider Threat Honeypot	Rilevano abusi di privilegi e accessi non autorizzati da utenti interni.	Visibilità su minacce interne.	Rischio di falsi positivi se non ben configurato.

Il limite principale degli honeypot tradizionali era la mancanza di scalabilità, automazione e specializzazione. Erano isole di inganno in un oceano di realtà, difficili da gestire, configurare e integrare con l'ecosistema di sicurezza esistente. Uno studio accademico del 2009 ha evidenziato come l'uso di *"advanced honeypots"* potesse risolvere i problemi di *throughput*, *latency* e sicurezza degli IDS tradizionali, riducendo significativamente i falsi allarmi[16].

2.4 L'Era Moderna: Le Distributed Deception Platforms (DDP)

Le moderne piattaforme di Deception superano questi limiti storici. Sono soluzioni centralizzate e automatizzate che creano un tessuto di inganno pervasivo e credibile, intrecciato con l'infrastruttura produttiva reale. Questo ecosistema è composto da tre elementi chiave:

- **Decoys (Trappole):** asset ad alta interazione (server, database, PLC, postazioni utente, dispositivi IoT) che appaiono come sistemi di produzione autentici. Possono emulare Active Directory, terminali vari, HMI industriali e molto altro.
- **Breadcrumbs (Briciole):** credenziali false, file di configurazione, cookie di sessione, cronologie di navigazione disseminati sugli asset reali per guidare gli attaccanti verso i decoys durante le loro attività di ricognizione e movimento laterale.
- **Lures (Esche):** asset a bassa interazione che emulano porte o servizi aperti per attirare l'attenzione iniziale dell'attaccante durante la fase di *scanning*.

Questa architettura distribuita trasforma l'intera rete in un ambiente ostile per l'attaccante, dove ogni mossa può portare alla scoperta. Come evidenziato dal framework MITRE Engage, "con l'adversary engagement, l'attaccante deve essere sbagliato solo una volta", ribaltando il paradigma tradizionale dove il difensore doveva avere sempre ragione[9].

2.5 ROI e Metriche per un Business Case

L'adozione della Deception Technology deve essere supportata da un solido business case, validato da studi, implementazioni reali e raccomandazioni governative:

- **Riduzione del Costo dei Breach:** la riduzione significativa del *dwell* time, da scale temporali dell'ordine dei mesi a finestre temporali molto più contenute, incide in modo diretto sulla diminuzione dei costi complessivi di gestione degli incidenti. Considerando un costo medio globale di un data *breach* pari a 4,44 milioni di dollari nel 2025, la contrazione dei tempi di permanenza dell'attaccante rappresenta uno dei principali fattori di mitigazione dell'impatto economico complessivo[1].

- **Efficienza del SOC:** la riduzione dei falsi positivi (fino al 95% secondo dati del vendor Acalvio[10]) permette ai gruppi di sicurezza di concentrarsi su minacce reali, aumentando drasticamente l'efficienza operativa e riducendo il *burnout* del personale.
- **Aumento del Costo per l'Attaccante:** studi accademici dimostrano che la Deception aumenta significativamente il tempo e le risorse necessarie per un attacco, agendo come deterrente economico[14].
- **Raccolta di Intelligence Unica:** il framework MITRE Engage evidenzia come le operazioni di *adversary engagement* consentano un incremento significativo della quantità e della qualità dell'intelligence raccolta rispetto ai modelli tradizionali, permettendo di ottenere decine di indicatori di compromissione per singola operazione invece di poche unità tipiche dei sistemi passivi[9].
- **Validazione Governativa:** il Dipartimento della Salute e dei Servizi Umani degli Stati Uniti (HHS) raccomanda ufficialmente l'uso di honeypot per *network intrusion detection*, particolarmente in ambienti critici e altamente regolamentati come il settore *healthcare*[17].

2.6 Considerazioni Operative e Limiti della Deception

Sebbene le Deception Technology offrano benefici rilevanti in termini di rilevamento precoce delle compromissioni, riduzione del *dwell time* e incremento della qualità degli *alert* di sicurezza, la loro adozione non può essere considerata un intervento puramente tecnologico né una soluzione autosufficiente. Tali capacità devono essere comprese e implementate come strumenti a supporto di una più ampia strategia di cybersecurity, coerente con il profilo di rischio dell'organizzazione, il contesto normativo di riferimento e il livello di maturità dei processi di sicurezza esistenti.

In assenza di una visione strategica chiara e di un modello di governo operativo strutturato, l'introduzione di tecnologie di Deception rischia di tradursi in un aumento della complessità infrastrutturale e operativa, senza generare i benefici attesi in termini di resilienza e capacità di risposta. La loro efficacia è infatti strettamente dipendente dalla qualità della progettazione, dall'allineamento con i processi di monitoraggio, *incident response e threat intelligence*, nonché dalla capacità dell'organizzazione di integrare tali strumenti all'interno di un modello di difesa multilivello.

Alla luce di tali considerazioni, è fondamentale valutare con attenzione le principali criticità operative associate all'adozione delle Deception Technology:

- **Costo Iniziale e Complessità:** le piattaforme di Deception *enterprise-grade* richiedono un investimento iniziale significativo e una configurazione complessa per garantire che gli asset ingannevoli siano credibili e ben integrati con l'infrastruttura reale.
- **Rischio di proliferazione di *decoy*:** una gestione non adeguatamente governata delle trappole può determinare una crescita incontrollata degli asset di deception, con conseguente aumento della complessità di manutenzione e del carico operativo, nonché il rischio che risorse non più correttamente aggiornate o presidiate diventino nel tempo potenziali punti di debolezza.
- **Rilevamento da Attaccanti Sofisticati:** attaccanti avanzati (APT) possono utilizzare tecniche di *fingerprinting* per identificare e *bypassare* gli asset di deception, specialmente se non sono mantenuti e aggiornati costantemente per riflettere i cambiamenti dell'ambiente produttivo.
- **Necessità di Aggiornamento Continuo:** per rimanere efficaci, le esche e le trappole devono evolvere insieme all'infrastruttura reale. Questo richiede un *effort* operativo continuo per garantire che il realismo degli asset ingannevoli non degradi nel tempo.

3. L'Apice della Difesa Proattiva: Automated Moving Target Defense (AMTD)

Se la Deception Technology ribalta l'asimmetria attaccante/difensore, l'Automated Moving Target Defense (AMTD) la porta a un livello esponenziale. L'AMTD è una strategia definita da Gartner come ad alto impatto strategico che rende la superficie d'attacco dinamica e imprevedibile, invalidando continuamente le informazioni raccolte dagli attaccanti e rendendo estremamente difficile la pianificazione e l'esecuzione di un attacco[4] [11].

3.1 Architettura e Meccanismi dell'AMTD

Una strategia AMTD si basa su motori di mutazione e diversità orchestrati automaticamente attraverso quattro layer principali:

1. **Mutation Engine:** implementa tecniche come *IP Shuffling* (rotazione continua degli indirizzi IP), *Port Hopping* (modifica dinamica delle porte dei servizi) e *Protocol Rotation* (alternanza dei protocolli di comunicazione).
2. **Diversity Engine:** gestisce *OS Polymorphism* (variazione periodica del sistema operativo o della versione del kernel), *Application Virtualization* (spostamento di container tra host fisici) e *Configuration Randomization* (modifica parametri di sistema).
3. **Orchestration Layer:** coordina le *policy* di mutazione, lo *scheduling* dei cambiamenti e l'integrazione con la piattaforma di *Deception*.
4. **Monitoring & Analytics:** misura metriche come *Attack Surface Reduction*, *Attacker Cost Increase* ed *Effectiveness Scoring*.

3.2 AMTD in Ambienti OT e Legacy: Sfide e Strategie di Mitigazione

L'integrazione dell'Automated Moving Target Defense (AMTD) in ambienti di **Tecnologia Operativa (OT)** e in sistemi **Legacy** rappresenta una delle sfide più significative per le organizzazioni che operano in settori critici come l'energia, la manifattura e i trasporti. Questi ambienti sono caratterizzati da un'estrema fragilità, requisiti di stabilità operativa non negoziabili e cicli di vita dei sistemi che si estendono per decenni.

Le Criticità dell'AMTD in OT/Legacy

I meccanismi di mutazione e diversità che sono il cuore dell'AMTD (come l'IP Shuffling, il Port Hopping o l'OS Polymorphism) sono spesso incompatibili con le esigenze di questi sistemi per i seguenti motivi:

1. **Requisiti di Stabilità e Tempo Reale:** i sistemi OT (SCADA, PLC, HMI) dipendono da comunicazioni a bassa latenza e indirizzi IP statici per garantire la sicurezza fisica e la continuità operativa. Una variazione dinamica può causare interruzioni, malfunzionamenti o, nel peggiore dei casi, situazioni di pericolo.
2. **Protocolli Proprietari e Sensibilità:** molti protocolli OT (es. Modbus, DNP3, Profinet) non sono progettati per tollerare interruzioni o cambiamenti di configurazione. Inoltre, i sistemi Legacy spesso non ricevono più aggiornamenti di sicurezza, rendendoli vulnerabili e incapaci di gestire i moderni meccanismi di difesa.
3. **Validazione e Certificazione:** qualsiasi modifica in un ambiente OT richiede una rigorosa e costosa ri-validazione per rispettare gli standard di sicurezza e le normative di settore. L'AMTD, per sua natura, introduce un cambiamento continuo, rendendo la validazione un processo quasi impossibile.

Strategie di Mitigazione e Difesa Ibrida

Per estendere i benefici della Sicurezza Attiva in queste zone sensibili, è necessario adottare un approccio ibrido e non intrusivo, che privilegi la **Deception passiva** rispetto alla mutazione attiva.

Strategia di Mitigazione	Descrizione	Beneficio per OT/Legacy
Segmentazione e Isolamento	Implementare una rigorosa segmentazione di rete (*Zero Trust*) tra l'ambiente IT e l'ambiente OT, limitando il traffico al minimo indispensabile.	Riduce drasticamente la superficie d'attacco e il rischio di movimento laterale dall'IT all'OT.
Deception Protocol-Aware	Utilizzare *Decoy* e *Honeypot* progettati per emulare specifici dispositivi e protocolli OT (es. un PLC Siemens o un server Modbus) senza interagire con l'asset reale.	Fornisce una capacità di rilevamento ad alta fedeltà (*near-zero false positives*) senza introdurre instabilità nel sistema produttivo.
AMTD sul Perimetro (Boundary)	Applicare le tecniche AMTD solo ai sistemi di frontiera (es. *jump servers*, DMZ, *firewall*) che fungono da ponte tra IT e OT.	Rende il punto di ingresso e il primo strato di movimento laterale imprevedibile, proteggendo indirettamente l'ambiente OT.
Honey Credentials e Breadcrumbs	Disseminare credenziali e file fittizi (*breadcrumbs*) sui sistemi IT che potrebbero essere utilizzati per accedere all'OT.	Cattura l'attaccante prima che raggiunga la rete OT, trasformando la fase di ricognizione in un evento di rilevamento.
Validazione Non Intrusiva	Prima di qualsiasi implementazione, eseguire test di non-interferenza rigorosi e validati dal team di ingegneria OT per garantire la stabilità operativa.	Assicura che l'introduzione di nuovi controlli non comprometta la sicurezza fisica o la continuità del *business*.

In sintesi, l'AMTD diretto sui sistemi OT è sconsigliato. La strategia vincente consiste nell'utilizzare la Deception Technology come meccanismo di rilevamento primario all'interno dell'OT e l'AMTD come meccanismo di frustrazione e deterrenza sul perimetro IT/OT, creando un "cuscinetto" dinamico che protegge i sistemi più sensibili. Questo approccio ibrido massimizza l'efficacia della Sicurezza Attiva mantenendo la stabilità operativa.

3.3 La Sinergia Perfetta: AMTD + Deception

L'integrazione di AMTD e Deception crea un meccanismo di difesa altamente efficace e costoso per l'attaccante, articolato in quattro fasi:

- **Frustrazione della Ricognizione:** l'AMTD rende l'ambiente di produzione instabile e caotico. Le scansioni di rete diventano obsolete in pochi minuti, costringendo l'attaccante a ri-scansionare continuamente con un costo elevato in termini di tempo e risorse.
- **Attrazione verso l'Inganno:** in questo ambiente dinamico, gli asset di Deception (Decoys) rimangono stabili e attraenti. L'attaccante, frustrato dall'instabilità, è psicologicamente spinto a interagire con gli unici asset che sembrano affidabili, cadendo così nella trappola.
- **Rilevamento Immediato:** non appena l'attaccante interagisce con un *Decoy*, viene generato un alert ad altissima fedeltà, permettendo un rilevamento quasi istantaneo.
- **Risposta e Adattamento:** l'alert può innescare una risposta automatizzata, come l'isolamento dell'asset compromesso o l'aumento della frequenza di mutazione dell'AMTD, rendendo ancora più difficile per l'attaccante continuare l'attacco.

4. AI e il Futuro della Difesa: Dal Gap di Velocità alla Difesa Autonoma

L'integrazione tra Intelligenza Artificiale (AI) e Deception Technology/MTD non è un *trend* futuro, ma una realtà consolidata nel 2024-2025. Questo è guidato anche dal concetto di *"Dual-Use AI"*: mentre gli attaccanti usano GenAI per attacchi sofisticati (il 16% degli incidenti nel 2025 coinvolgerà *AI-offensive*), i difensori rispondono con *"deceptive AI"* che aumenta l'efficacia della Deception del 30%. L'AI sta trasformando queste tecnologie da strumenti reattivi a sistemi proattivi e predittivi, automatizzando e potenziando ogni fase del ciclo di vita della difesa.

4.1 Il "Velocity Gap" e l'Urgenza della Difesa Preemptive

Il recente caso di una campagna di cyber-spionaggio orchestrata da AI, documentata da Anthropic, ha evidenziato una profonda vulnerabilità nel modello di sicurezza tradizionale: il **"velocity gap"**[23]. In questa campagna, gli agenti AI hanno eseguito l'80% o più delle azioni di attacco, operando a una velocità e una scala inimmaginabili per un team umano.

> *"Il tempo richiesto per un analista umano per rivedere e tracciare l'attacco e intraprendere un'azione significativa in risposta non è più sufficiente per tenere il passo con il ciclo di vita del cyberattacco."*[23]

Questa velocità di esecuzione annulla l'efficacia delle strategie tradizionali di Detect and Respond (D&R). La priorità strategica si sposta quindi sulla cybersecurity preemptive, che mira a interrompere le catene di attacco autonome prima che possano causare danni. Gartner raccomanda di dare priorità a tecniche come l'AMTD e l'Advanced Cyber Deception per difendersi proattivamente contro minacce sofisticate e mirate[23].

4.2 I Nuovi Principi Architetturali: Denial, Disruption, Deception

Per colmare il *velocity gap*, i *leader* di prodotto devono iniziare immediatamente a concentrarsi sui principi di negazione (*denial*), interruzione (*disruption*) e inganno (*deception*), considerandoli come pilastri architettonici non negoziabili per interrompere preventivamente le catene di attacco autonome[23].

Principio	Descrizione	Meccanismo Tecnologico
Denial (Negazione)	Impedire all'attaccante di ottenere informazioni affidabili sull'ambiente.	AMTD (*IP Shuffling, OS Polymorphism*)
Disruption (Interruzione)	Interrompere attivamente il movimento laterale e la persistenza dell'attaccante.	Isolamento automatico degli *host*, revoca dinamica delle credenziali.
Deception (Inganno)	Catturare l'attaccante e raccogliere *intelligence* ad alta fedeltà.	*Distributed Deception Platforms (Decoys, Breadcrumbs, Lures).*

4.3 LLM-Powered Honeypots: La Rivoluzione dell'Interazione

I Large Language Models (LLM) stanno trasformando radicalmente gli honeypot da sistemi statici a entità intelligenti e interattive. Un paper del 2024 dimostra come gli LLM possano creare honeypot capaci di engagement sofisticato con attaccanti avanzati, rendendo quasi impossibile distinguerli da sistemi reali. Progetti open-source come HoneyLLM e shelLM stanno già creando shell dinamiche basate su LLM, neutralizzando le tecniche di fingerprinting che tradizionalmente permettevano agli attaccanti di identificare ed evitare le trappole.

La Nostra Implementazione Operativa

Anche noi, con il team AI4Cyber, abbiamo sviluppato internamente un framework per honeypot ad alta interazione basati su Large Language Models eseguiti in locale – un requisito fondamentale per ambienti governativi dove i dati non possono uscire dal perimetro.

L'architettura si articola su tre moduli:

- **Terminal Protocol Proxy:** gestisce le connessioni SSH/Telnet in ingresso, presentando all'attaccante un'interfaccia indistinguibile da un sistema reale.
- **Prompt Manager:** orchestra il contesto conversazionale, integrando algoritmi di pruning per gestire i limiti della context window e un sistema di classificazione dei comandi che distingue tra operazioni sicure da emulare e tentativi di evasione.
- **LLM Engine:** il cuore del sistema, testato con modelli open-source (Mistral, Gemma2, Phi4, Llama3.1, Qwen2.5) e proprietari (GPT-4, Claude), capace di generare risposte contestuali che simulano in modo credibile un terminale Linux, filesystem incluso.

I test condotti hanno confermato la capacità del sistema di mantenere l'engagement con attaccanti reali per sessioni prolungate, raccogliendo tattiche, tecniche e procedure degli attaccanti (TTP) e indicatori di compromissione (IOC) di valore operativo. A differenza degli honeypot tradizionali, dove l'attaccante può identificare la trappola con tecniche di fingerprinting, nel nostro framework ogni sessione è unica e ogni risposta è generata dinamicamente – rendendo il decoy virtualmente indistinguibile da un sistema di produzione[22].

4.4 AI-Native Deception Platforms: Automazione End-to-End

Le piattaforme commerciali hanno integrato l'AI in ogni fase del ciclo di vita della deception:

- **Pre-Attack Intelligence:** grazie all'analisi automatizzata dell'attack surface di Active Directory (150+ vettori), l'AI identifica proattivamente account ad alto rischio (Kerberoastable, Shadow Admins) e misconfigurazioni critiche, trasformando vulnerabilità latenti in rischi immediatamente osservabili.
- **Automated Design:** un *recommendation engine* ML-powered genera automaticamente oltre 100 attributi AD per i *honeytoken*, colmando lo *skill gap* e garantendo il massimo realismo.
- **Intelligent Triaging:** un *threat analytics engine* basato su ML reduce centinaia di eventi a pochi *alert* azionabili, correlandoli con l'EDR e mappandoli automaticamente al *framework* MITRE ATT&CK.

- **Generative AI per Content:** gli LLM generano contenuti *context-aware* per i *decoys* ad alta interazione, allineati al *vertical* dell'organizzazione per massimizzare la credibilità.

4.5 AI per Automated Moving Target Defense (AMTD) e Sistemi Cyber-Immuni

L'AI sta rendendo l'AMTD ancora più dinamica e intelligente, evolvendo verso sistemi di *"Self-Healing"*. Progetti come ADA (*Adaptive Defense Agent*) utilizzano l'AI per l'adattamento in tempo reale delle difese, mentre la ricerca accademica si concentra sull'uso dell'AI per l'*automated decision-making* in MTD, ottimizzando dinamicamente i meccanismi di difesa, specialmente in ambienti IoT.

Questa evoluzione è in linea con la previsione di Gartner secondo cui, entro il 2030, il 75% o più delle grandi imprese avrà implementato capacità di sistema cyber-immune autonome contro le minacce guidate dall'AI[23].

4.6 Mercato in Esplosione e Implicazioni Strategiche

Il mercato degli *AI Deception Tools* è previsto crescere da $640.4 milioni nel 2024 a $6.4 miliardi nel 2033, con un CAGR del 28%. Gartner stima, inoltre, un ritorno sull'investimento (ROI) di 2-3x in soli 18 mesi per le tecnologie di sicurezza *preemptive* basate su AI. Questa crescita esponenziale valida la tesi che l'AI-powered deception è il futuro della cybersecurity proattiva. Per i C-Level, questo significa:

- **ROI Amplificato:** l'automazione end-to-end riduce drasticamente i costi operativi e lo skill gap.
- **Vantaggio Competitivo Asimmetrico:** l'AI crea un ambiente dove il costo per l'attaccante è massimizzato e il costo per il difensore è minimizzato.
- **Allineamento con le Previsioni di Mercato:** l'investimento in *AI-powered deception* è allineato con la previsione di Gartner che il 50% della spesa in IT security sarà su tecnologie *preemptive* entro il 2030.

5. Visione Strategica e Piano di Integrazione

Un percorso di adozione corretto prevede una sequenza ben definita: prima la diffusione pervasiva della Deception sull'intera infrastruttura, successivamente l'introduzione graduale di meccanismi di Moving Target Defense su asset selezionati e, solo in una fase avanzata, l'automazione adattiva guidata dall'intelligenza artificiale. Questo consente di ottenere benefici immediati in termini di *detection* riducendo al minimo i rischi di instabilità e costruendo nel tempo una traiettoria di maturità sostenibile.

L'adozione di Deception e MTD non è un semplice acquisto di tecnologia, ma una trasformazione strategica del modello di sicurezza. Richiede una visione chiara e un piano di implementazione graduale.

5.1 Modello di Maturità della Sicurezza

L'integrazione di queste tecnologie permette di evolvere il modello di sicurezza attraverso tre livelli di maturità:

- **Difesa Passiva (Livello 1):** l'organizzazione si affida a difese perimetrali e risponde agli incidenti dopo che si sono verificati. È un approccio reattivo e inefficiente.
- **Difesa Proattiva (Livello 2):** l'organizzazione implementa Deception e MTD per cacciare attivamente gli attaccanti all'interno della rete, riducendo drasticamente il *Dwell Time* e i danni.
- **Difesa Predittiva (Livello 3):** utilizzo di AI e *Machine Learning* sui dati raccolti (specialmente dalle piattaforme di Deception) per anticipare le mosse degli attaccanti e adattare dinamicamente le difese attraverso AMTD. Questo livello rappresenta la frontiera della cybersecurity autonoma.

5.2 Integrazione nel Security Operations Center (SOC)

Deception e MTD non sono silos tecnologici, ma catalizzatori che potenziano e trasformano i processi esistenti del SOC. L'adozione di queste tecnologie non è solo un acquisto di strumenti, ma un investimento strategico sul capitale umano. Richiede un *upgrade* qualitativo delle competenze degli analisti, che passano dalla gestione meccanica di migliaia di falsi allarmi ripetitivi alla gestione di pochi incidenti complessi ma reali. Questo permette di valorizzare il *team* di sicurezza, trasformandolo da un centro di costo reattivo a un centro di intelligence proattivo.

- **Incident Response (IR):** gli *alert* di Deception sono eventi ad altissima fedeltà (*Falsi positivi strutturalmente minimi / near-zero false positives*). Questo permette di attivare *playbook* di risposta automatizzati (es. isolamento dell'host, blocco dell'account, quarantena della sessione) con la massima fiducia, riducendo drasticamente il *Mean Time to Respond* (MTTR).
- **Threat Hunting:** i gruppi di *threat hunting* possono usare i dati di Deception per avviare indagini mirate e proattive. Se un attaccante usa una certa credenziale esca (*honey credential*), dove altro ha provato a usarla? L'esca diventa un filo da seguire per svelare l'intera campagna di attacco.
- **Threat Intelligence:** i *Decoy* ad alta interazione sono mini-laboratori forensi isolati. Permettono di raccogliere TTPs, campioni di *malware* e indicatori di compromissione (IOC) specifici per la propria organizzazione, creando un *feed* di *intelligence* proprietario e di valore inestimabile per rafforzare le difese.

5.3 Roadmap di Adozione Strategica: Crawl, Walk, Run

L'implementazione dovrebbe seguire un approccio graduale e pragmatico per garantire il successo, minimizzare i rischi e massimizzare il ROI.

Fase	Durata	Azioni Chiave	Obiettivo	KPI di Successo
Crawl	**3–6 mesi**	*Pilot* Deception: implementare una piattaforma di Deception in un segmento di rete critico ma controllato (es. una VLAN con dati sensibili). Creare esche e trappole di base. Integrare gli *alert* nel SIEM.	Rilevare immediatamente compromissioni interne, ridurre i falsi positivi.	Rilevamento di 1–2 incidenti reali, riduzione Falsi Positivi >50%.
Walk	**6–12 mesi**	*Scale* Deception & *Pilot* MTD: Estendere la Deception a tutta l'infrastruttura (inclusi *cloud* e OT*). Avviare un *pilot* di MTD (es. *IP shuffling*) su asset non critici. Iniziare a usare i dati per il *threat hunting*.	Ottenere visibilità completa, aumentare i costi per l'attaccante, iniziare a costruire una difesa proattiva.	Copertura >80% asset, >20 incidenti rilevati, riduzione *Dwell Time* >50%.
Run	**12+ mesi**	*Automate & Integrate* (AMTD): Implementare l'AMTD. Integrare Deception e AMTD in modo che si alimentino a vicenda. Automatizzare i *playbook* di risposta basati sugli *alert* di Deception.	Raggiungere un modello di *Active Defense* maturo, con capacità di risposta autonoma e resilienza dinamica.	*Dwell Time* <24 ore, risposta automatizzata >90%, ROI positivo.

**con protocolli di sicurezza passiva per ambienti legacy e previa validazione rigorosa di non-interferenza operativa.*

5.4 Metriche di Successo e KPI per il Board

Per giustificare l'investimento e misurare il successo del programma, è fondamentale presentare al *board* metriche chiare e orientate al *business*:

- **Riduzione del Dwell Time:** da >200 giorni a <24 ore (obiettivo: riduzione del 95%).
- **Mean Time to Detect (MTTD):** da mesi a minuti/secondi per gli *alert* di Deception.

- **Riduzione dei Falsi Positivi:** percentuale di riduzione degli *alert* non azionabili nel SIEM.
- **Numero di Incidenti Rilevati:** numero di incidenti reali identificati esclusivamente tramite Deception.
- **ROI del Programma:** calcolato come (Costo evitato dei *breach* – Costo della soluzione) / Costo della soluzione.
- **Aumento del Costo per l'Attaccante:** metrica qualitativa basata sull'analisi dell'*intelligence* raccolta.

6. Deception Technologies e NIS2: Dalla Compliance alla Cyber-Resilience Attiva

6.1 NIS2: Un Cambio di Paradigma nella Governance Europea

La Direttiva UE 2022/2555 (NIS2) segna il passaggio definitivo da una sicurezza basata su checklist a una strategia fondata sulla **resilienza operativa** e sulla **gestione proattiva del rischio**. Per le organizzazioni classificate come *Essenziali e Importanti*, la conformità non è più un traguardo statico, ma un obbligo continuo di adottare misure "adeguate e proporzionate" allo stato dell'arte tecnologico (Art. 21).

In questo scenario, la Deception Technology e l'Automated Moving Target Defense (AMTD) non sono semplici opzioni tattiche, ma pilastri strutturali per soddisfare i requisiti di **rilevamento tempestivo e accountability del management**.

6.2 Mappatura Strategica: Deception/AMTD vs Requisiti NIS2

L'adozione di tecnologie di Sicurezza Attiva risponde direttamente ai principali obblighi sanciti dall'Articolo 21 della Direttiva:

Requisito NIS2 (Art. 21)	Contributo della Deception & AMTD	Valore Strategico per la Compliance
Gestione del Rischio (21.1)	Introduzione di asset ingannevoli (server, identità, API, OT) per intercettare ricognizione e *lateral movement*.	Trasforma il rischio da teorico a **osservabile e misurabile**, riducendo l'impatto economico tramite la contrazione del *dwell time*.
Incident Detection & Response (21.2 b–c)	Generazione di *alert near-zero false positives*. Ogni interazione con un *decoy* è un evento ad alta confidenza.	Abilita **playbook automatici** di isolamento e risposta, garantendo la tempestività richiesta dal legislatore.
Supply Chain Security (21.2 d)	Deploy di *honey credentials* e API token fittizi in ambienti condivisi con partner e fornitori.	Rende monitorabile l'area più critica della NIS2: la **supply chain digitale**, rilevando abusi di account di terze parti.
Efficacia delle Misure (21.2 f)	Fornisce KPI oggettivi: riduzione MTTD/MTTR, numero di incidenti reali intercettati, costo per l'attaccante.	Offre evidenze concrete per **audit e ispezioni** delle autorità competenti, dimostrando l'efficacia reale delle difese.

6.3 La Deception come "Misura Proporzionata"

La NIS2 introduce il concetto di **proporzionalità**. Per settori critici come Sanità, Energia, Trasporti e Finanza, l'assenza di capacità di rilevamento avanzato può essere configurata come un gap di diligenza tecnica. Alla luce del panorama delle minacce (APT, Ransomware mirato, Insider Threat), la Deception rappresenta la risposta tecnologicamente proporzionata per proteggere infrastrutture complesse e ambienti cloud/ibridi.

6.4 Governance e Accountability (Art. 20)

La Direttiva pone la responsabilità della cybersecurity direttamente in capo agli organi direttivi. La Deception Technology supporta questa responsabilità trasformando la sicurezza in uno strumento di **intelligence strategica**:

- **Reporting basato su dati reali:** il Board riceve analisi su superfici di attacco effettivamente sfruttate, non solo su vulnerabilità teoriche.
- **Decision-making informato:** la visibilità sulle TTP (Tactics, Techniques, and Procedures) degli avversari permette investimenti mirati e una governance basata sull'evidenza.

6.5 Verso la Cyber-Resilience Attiva

La NIS2 non è una direttiva puramente difensiva; è il catalizzatore per una **resilienza sistemica**. L'integrazione di Deception, AMTD e AI-powered security costituisce un modello di **Active Defense** che supera il concetto di "protezione del perimetro". In linea con l'evoluzione normativa europea, la sicurezza moderna è definita dalla capacità di un'organizzazione di **resistere, adattarsi e recuperare** mentre è sotto attacco, trasformando l'infrastruttura stessa in un'arma di difesa.

7. Conclusione

L'adozione di una strategia di Sicurezza Attiva basata su Deception Technology e Automated Moving Target Defense non è più una scelta tecnologica, ma un **imperativo strategico** per qualsiasi organizzazione che voglia sopravvivere e prosperare nel panorama delle minacce moderne. Questo approccio permette di passare da una postura reattiva, costosa e inefficiente a una proattiva, intelligente e resiliente.

Ribaltando l'asimmetria storica tra attaccante e difensore, queste tecnologie non solo riducono drasticamente il rischio e il costo dei *data breach*, ma trasformano la sicurezza da un centro di costo a un centro di *intelligence* strategica, fornendo una visibilità senza precedenti sulle minacce reali che colpiscono l'organizzazione. Inoltre, l'allineamento con i requisiti della **Direttiva NIS2** garantisce non solo la conformità normativa, ma una reale capacità di difesa e resilienza operativa.

L'investimento in Deception e MTD, specialmente se potenziato dall'Intelligenza Artificiale, rappresenta il passo più efficace che un'organizzazione possa compiere oggi per costruire una difesa a prova di futuro, proteggere i propri asset critici e garantire la continuità del *business*. Si entra in una nuova era di conflitto cibernetico automatizzato, dove l'AI difensiva autonoma rappresenta la risposta più efficace contro le minacce *AI-driven*.

Bibliografia

[1] IBM. (2025). *Cost of a Data Breach Report 2025*. IBM Security. (Accessed: December 2025)

[2] Gartner. (2025). *Top Strategic Technology Trends for 2025*. Gartner, Inc. (Accessed: December 2025)

[3] Forrester. (2024). *The Forrester Wave™: Deception Technology, Q4 2024*. Forrester Research. (Accessed: December 2025)

[4] Morphisec. (2024). *Moving Target Defense: Your Complete Guide*. Morphisec. (Accessed: December 2025)

[5] National Bureau of Standards. (1976). *Glossary for Computer Systems Security*. Federal Information Processing Standards Publication 39 (FIPS PUB 39). (Accessed: December 2025)

[6] Stoll, C. (1989). *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage*. Doubleday. (Accessed: December 2025)

[7] Cheswick, B. (1992). *An Evening with Berferd, in Which a Cracker is Lured, Endured, and Studied*. AT&T Bell Laboratories. (Accessed: December 2025)

[8] CounterCraft. (2024). *Deception and Automated Moving Target Defense*. CounterCraft. (Accessed: December 2025)

[9] MITRE. (2024). *MITRE Engage Framework*. The MITRE Corporation. (Accessed: December 2025)

[10] Acalvio. (2023). *The Business Value of Deception Technology*. Acalvio Technologies. (Accessed: December 2025)

[11] Gartner. (2024). *Emerging Tech: Security – Automated Moving Target Defense*. Gartner, Inc. (Accessed: December 2025)

[12] Grand View Research. (2024). *Deception Technology Market Size, Share & Trends Analysis Report*. Grand View Research. (Accessed: December 2025)

[13] MarketsandMarkets. (2024). *Moving Target Defense Market – Global Forecast to 2030*. MarketsandMarkets. (Accessed: December 2025)

[14] Ferguson-Walter, K., et al. (2021). *Does Cyber Deception Improve Defense?*. USENIX Security Symposium. (Accessed: December 2025)

[15] NIST. (2024). *Cybersecurity Framework 2.0*. National Institute of Standards and Technology. (Accessed: December 2025)

[16] Singh, G., & Ramanujam, S. (2009). *Intrusion Detection System Using Advanced Honeypots*. International Journal of Computer Science and Information Security, 5(1). (Accessed: December 2025)

[17] U.S. Department of Health & Human Services. (2018). *Using Honeypots for Network Intrusion Detection*. HHS Cybersecurity Program. (Accessed: December 2025)

[18] Ng, C. K., Pan, L., & Xiang, Y. (2018). *Honeypot Frameworks and their Applications: A New Framework*. Springer Briefs on Cyber Security Systems and Networks. Springer Nature Singapore. ISBN 978-981-10-7738-8. (Accessed: December 2025)

[19] Acalvio Technologies. (2023). *Definitive Guide to Deception 2.0: Cybersecurity Manual for Distributed Deception Solutions*. Foreword by Dr. Gerhard Eschelbeck. Acalvio Technologies. (Accessed: December 2025)

[20] Sanders, C. *Intrusion Detection Honeypots: Detection through Deception*. Applied Network Defense. (Accessed: December 2025)

[21] AWS Security Blog. (2023). *How AWS threat intelligence deters threat actors*. Amazon Web Services. (Accessed: December 2025)

[22] Tinexta Defence. (2025). *AI4Cyber: Applicazione di Modelli Linguistici a Tecnologie di Deception ad Alta Interazione*. AI4Cyber Studio. (Accessed: December 2025)

[23] Gartner. (2025). *Emerging Tech: AI Vendor Race – AI Espionage Campaign Emphasizes Need for Preemptive Cybersecurity*. ID G00844051. (Accessed: December 2025)

tinexta
defence

Next | Donexit | Foramil | Innodesi

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

#TinextaDefenceBusiness