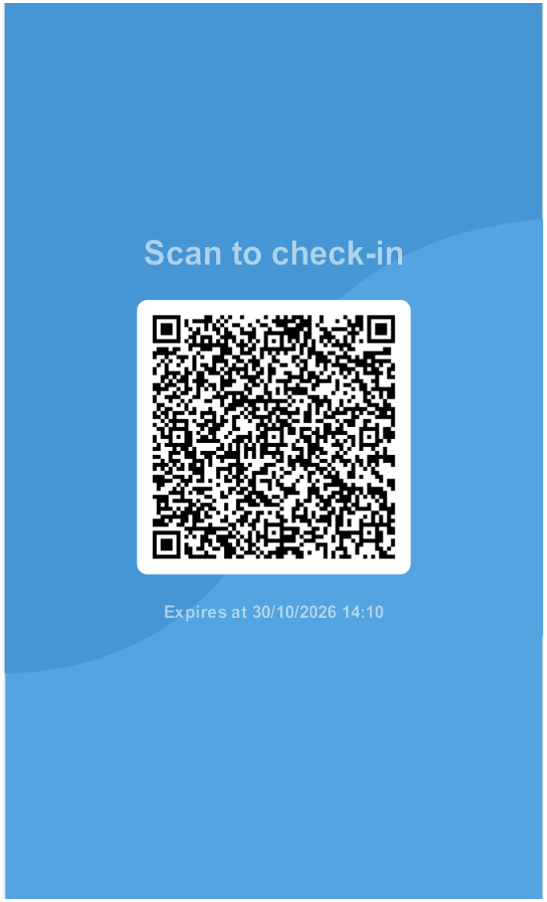
SCHOOL OF LECT

Security Engineering - Lecture

E:SX6VMP6XE3-00001

- Friday, 30 October 2026
- 13:10 - 14:00

Seminar Room 2



Security Engineering

INFR11208 (UG4) // INFR11228 (MSc)



Daniel W. Woods
Email: daniel.woods@ed.ac.uk

Theme of the Lecture:
Blocking, detecting and responding trades off the risk of compromise against the risk of disruption.

Perimeters are leaky, now what?

Attack Chain

1. **Get into network:** social engineering, malicious website, fraudulent email, stolen creds, vulnerable systems
   --> We can't reliably win here.
2. **Run something:** malware or built-in tools
3. **Get more access:** steal credentials, escalate privileges, exploit other devices
4. **Get what they came for:** exfiltrate data, encrypt systems, manipulate processes etc

So we block, detect and respond at Stages 2—4.

Simplified from the Lockheed Martin Cyber Kill Chain and MITRE ATT&CK.

Blocking: limiting what attackers can do and reach

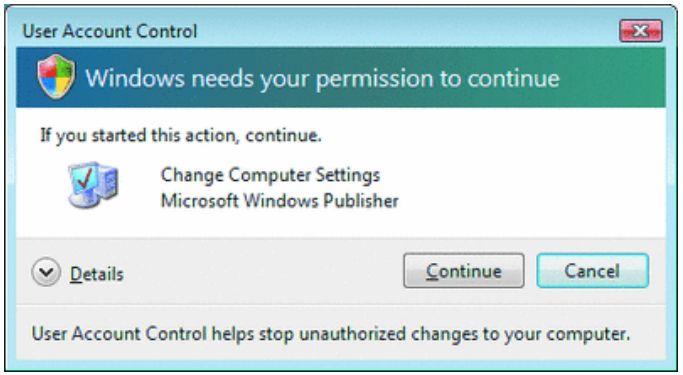
Preventing 1 click install

~~Security~~ Engineering in 1990s

- ~~Threat~~ Business model
  - Build the OS as a platform
  - Let 3rd party developers build value
    - 2-sided market
  - Make installing software easy
    - User runs in root
    - Apps run as root
- Security mechanisms....
  - Hope user only connects with trusted parties

What changed and when?

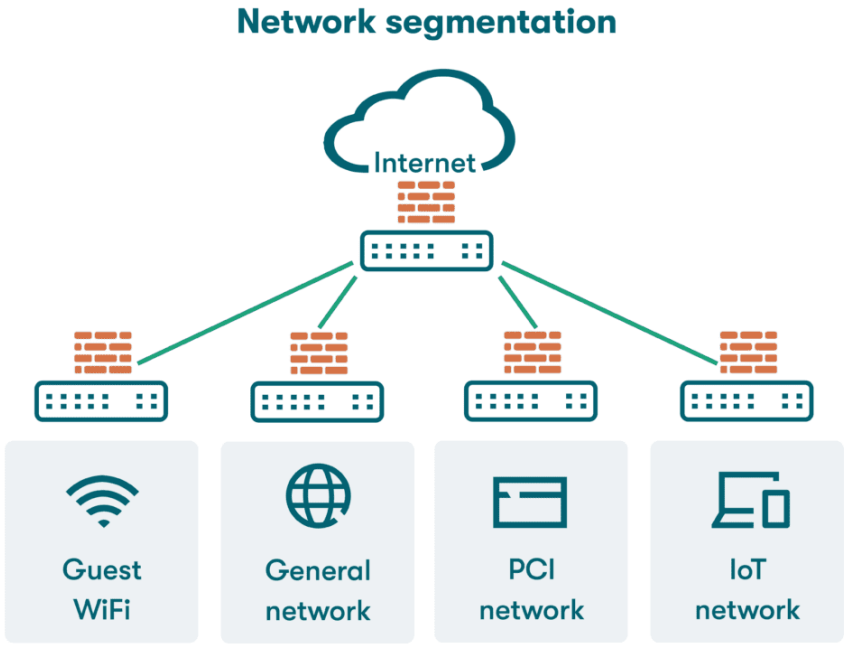
- Attackers started sharing malware via email and web
- Windows UAC (Vista, 2007), run as user by default
- Gatekeeper in MacOS (from 2012) checks if software signed by a dev
- Problem is users continue to click through
  - And get angry if you don't let them



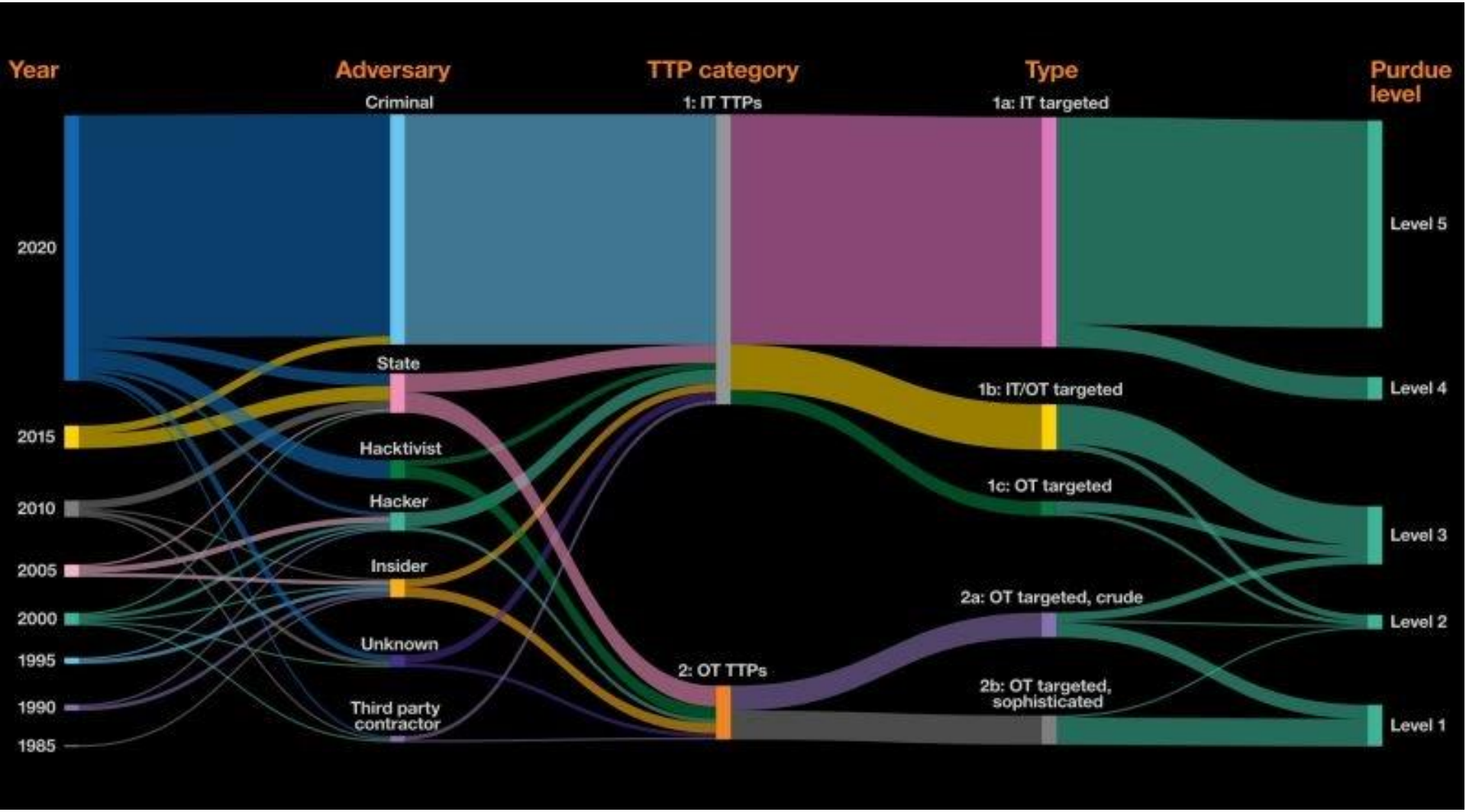


Network Segmentation

- Air gaps
  - The gold standard
    - But hard to achieve
    - See Stuxnet, infected USB drive
  - Hard for IT to administer
    - See backups
- Internal firewalls (+ monitoring)
  - Less guarantees, more usable
- IT – OT separation
  - Especially important for CNI



Capabilities and segmentation



Source: https://www.linkedin.com/feed/update/activity:7349028145104142336

Attack surface reduction

- **Deny Specifics:**
  Remove or add friction to functionality used by attackers but not legitimate users
  - USBs
  - Excel Macros --> send "salaryplan.xlsx"
  - Command line tools for most users
- **Allow specifics:**
  Maintain list of permitted apps, blocking everything else
  - At the cost of user friction
  - Maybe easier in SaaS age?

Microsoft says RIP Windows XP AutoRun for USB

By Liam Tung
Feb 10 2011 6:00AM

The feature that became an attack vector.

Learn > Microsoft 365 >

Macros from the internet are blocked by default in Office

Key control category	Signal strength
Hardening techniques System configuration management tools, such as an active directory group policy, which enforce and redeploy configuration settings to systems	5.58
Privileged access management Managing desktop or local administrator privileges via endpoint privilege management (EPM)	2.92
Endpoint detection and response Operating advanced endpoint security	2.23
Logging and monitoring Operating a security operations center (SOC) and/or having an outsourced managed security service provider (MSSP) with the following capabilities at a minimum: a. Established incident alert thresholds b. Security incident and event management (SIEM) monitoring and alerting for unauthorized access connections, devices and software	2.19
Patched systems Patching common vulnerability scoring system (CVSS) v3 high severity 7.0-8.9 vulnerabilities across the enterprise within 7 calendar days of release	2.19

https://www.guycarp.com/content/dam/guycarp-rebrand/pdf/Insights/2023/Using_data_to_prioritize_cybersecurity_investments_report.pdf

Managing Admin Privileges

Threat Model vs Business Reality

- Ransomware in 2011 vs 2021
  - Encrypt your machine vs network
- Sys admins love Active Directory + uniform local admin passwords

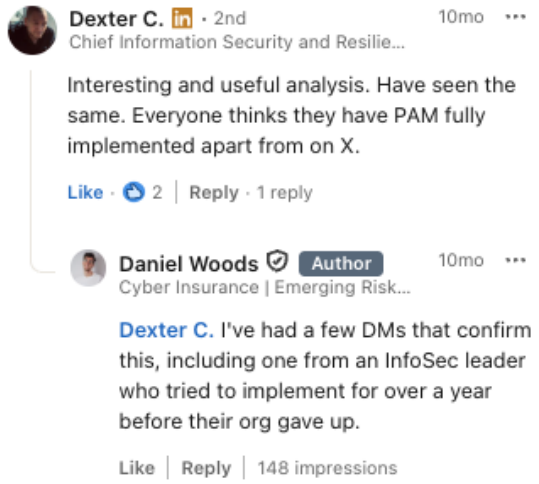
Security Mechanism

- Microsoft LAPS sets unique local admin password
  - Least privilege on machines
- Privileged Access Management
  - Store admin creds in central vault
  - Just in time access, expiring creds
  - Monitoring (logs + review)
- Sys admins hate it

[My Article]



"More than 66% of organizations compromised by a ransomware attack saw ransomware deployment linked to the compromise of a highly privileged account on their Active Directory environment."



Zero trust as privilege reduction

ZT: Don't grant access based on network location alone, instead verify identity and device on every request.

- Partly about improving confidence in authentication
  - Phishing resistant MFA
  - Device trust
- Partly about monitoring
  - Evaluate session risk (e.g. impossible travel)
- But mostly about limiting privilege
  - Reduce the time horizon of auth tokens
    - Auth more often
  - Reduce the breadth of auth tokens
    - Only let auth apply to specific services

Blocking as a people problem

- Trading off security vs user desires
  - Adding friction to install software and access data
  - Disabling rarely used functionality
  - Segmenting systems
- Harder to quell IT admin desires
  - Want seamless fixes across the entire network (see Active Directory)
    - Convenient for admins and ransomware
  - Don't want to review and approve
    - Biggest barrier to allowlist?
  - Complain much more loudly and have operational control
  - Security as sociology
- Visible and vocal user friction vs invisible but highly effective controls
  - Detection and response as the fix?

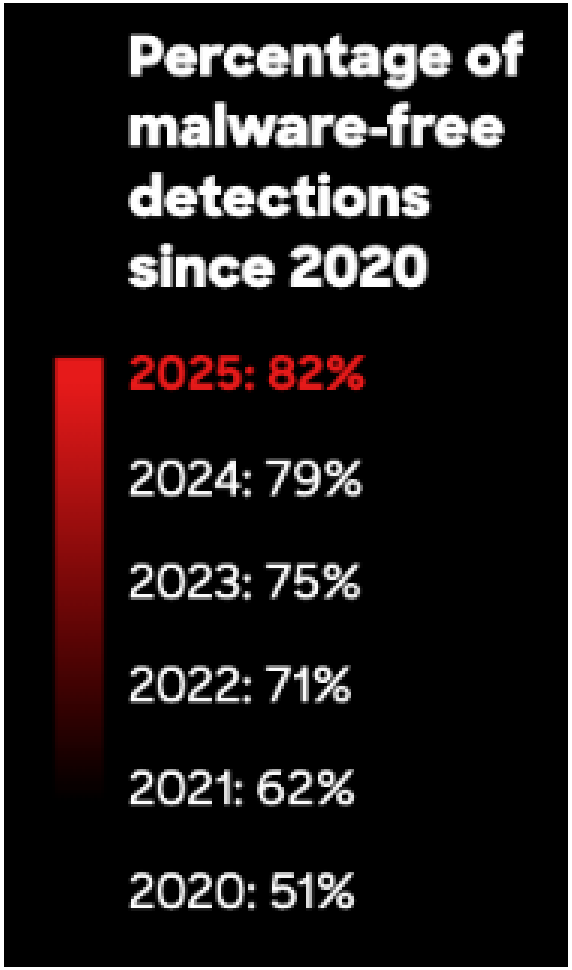


Detection and Response: Find suspicious activity

Post-compromise actions

- Malware (encoded instructions)
  - Viruses 1980s
    - Infect existing files, spread via USB/floppy
  - Worms and trojans (1990s and 2000s)
    - Worms autonomous via vulnerable and misconfigured devices
    - Trojans largely via email attachments
  - Hands on keyboard malware
    - Attacker directly installs corporate ransomware
- Malware free (live instructions)
  - Living off the land
  - Credential abuse
  - Installing RMMs

Perversely, more manual over time



Share of CrowdStrike detections not using malware

Detecting Business Email Compromise

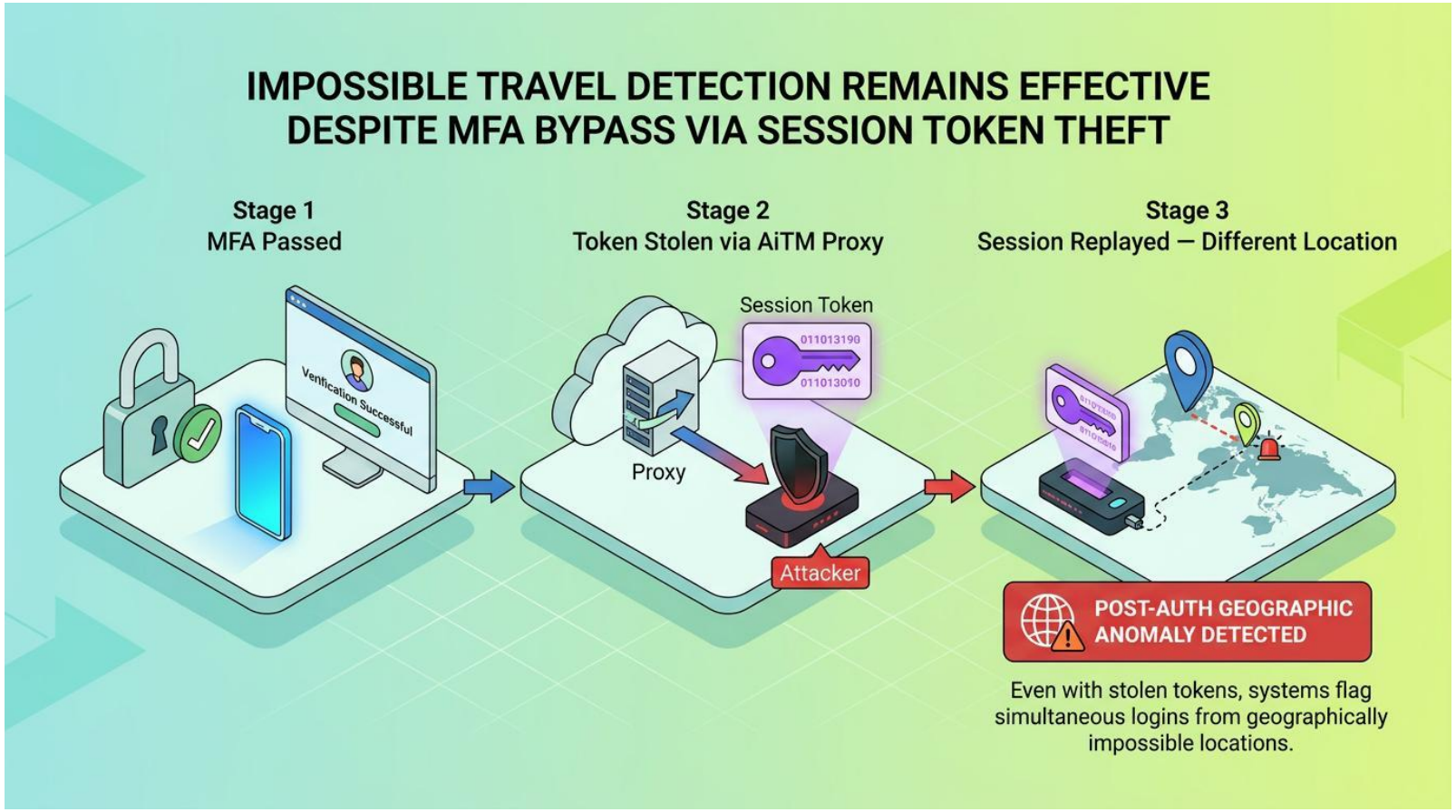
1. Login from new IP address
2. Accesses some emails
3. Changes some settings
4. …

Usually malware-free, with attacker taking actions.

Detecting Business Email Compromise

1. Login from new IP address
   - Is that IP known for maliciousness
   - Is the geographic location consistent with user
   - Is the time consistent with user
   - What device/user agent?
2. Accesses some emails
   - Does it look normal
3. Changes some settings
   - Create inbox redirects
4. ...

Potential detection signals 1



Attacker's network location shows up in your logs, especially if selling mass market access as a service.

Source: https://abnormal.ai/learning/impossible-travel-cybersecurity

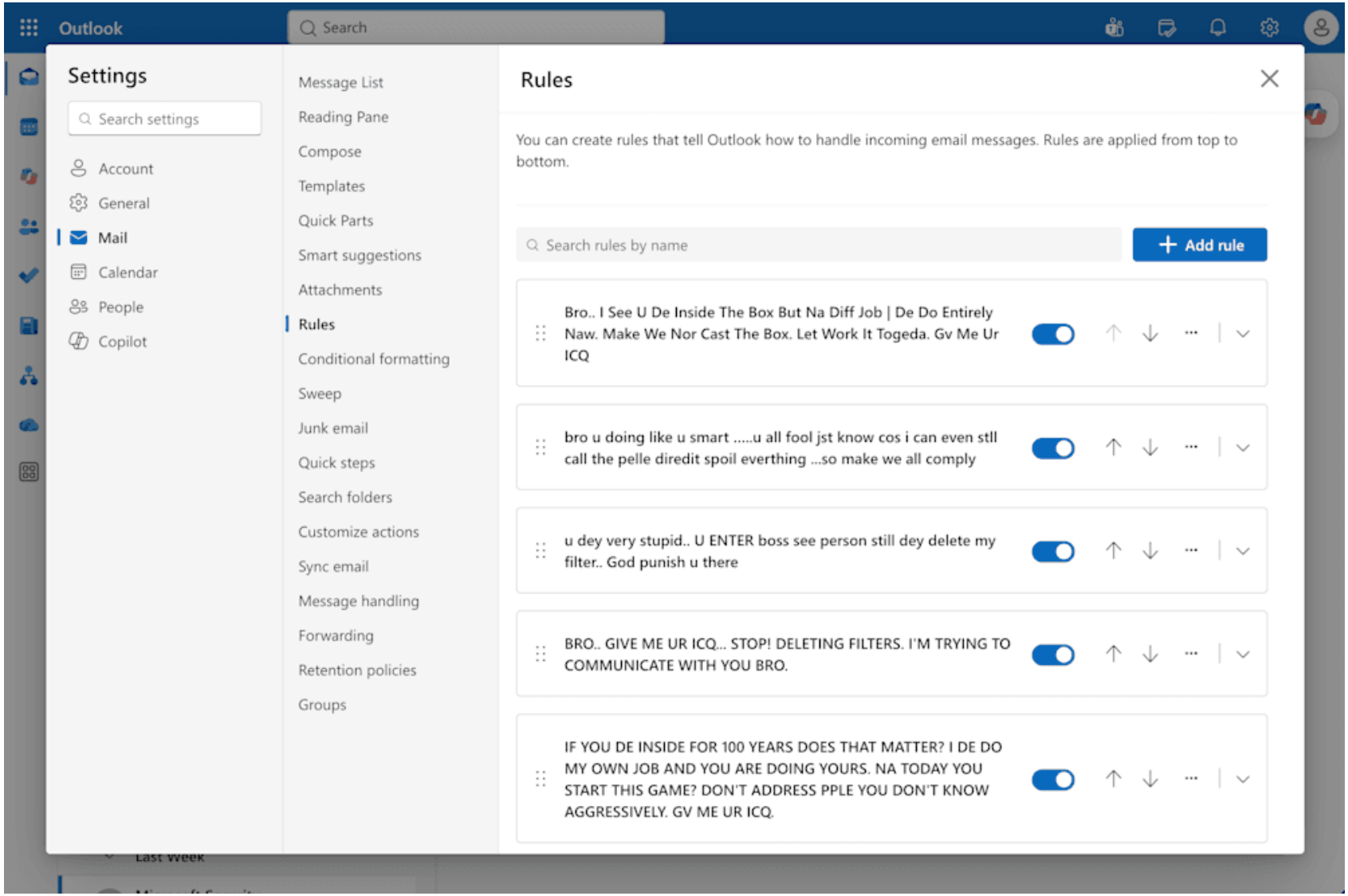
Potential detection signals 2

- Real world server farms or infected devices
- Criminal business model to sell access to "residential proxies"
- Detection idea is to maintain a list of these IP addresses



https://www.justice.gov/opa/pr/arizona-woman-sentenced-17m-information-technology-worker-fraud-scheme-generated-revenue

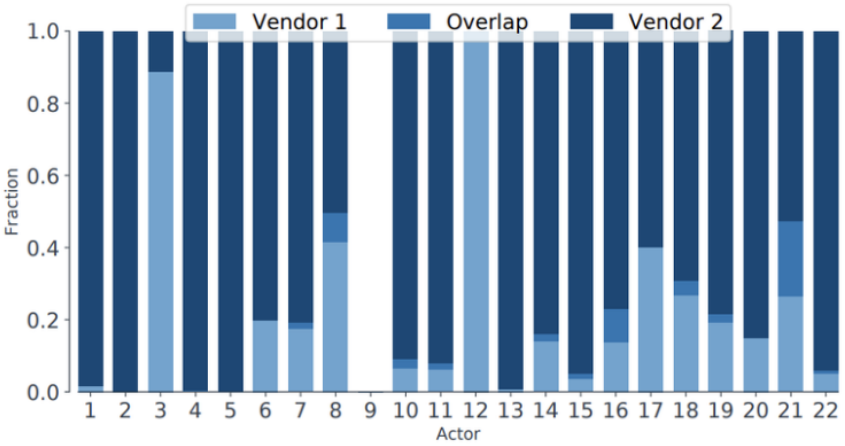
Potential detection signals 3



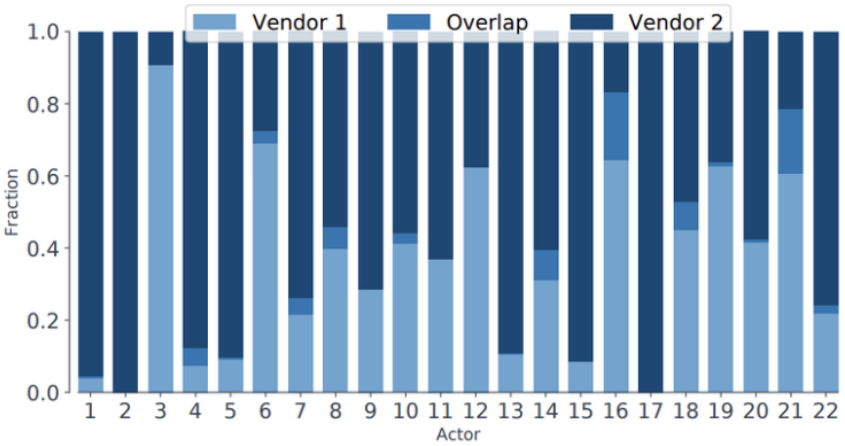
https://www.coalitioninc.com/en-gb/blog/wirespeed/inbox-in-fighting-a-window-into-bec-subculture

Why detection signals fail

- Legitimate reasons for "impossible" travel
  - VPNs, mobile carriers, gateways, inaccurate tags on IP addresses, delegated access…
- Legitimate reasons for "malicious" IP addresses
  - Shared infra (cloud, residential proxies, VPN exit nodes etc)
  - The threat intel provider was wrong (stale data, erroneous honey pot data etc)
- Legitimate reasons for mailbox rules



(a) IP indicators, avg. overlap 2,8%



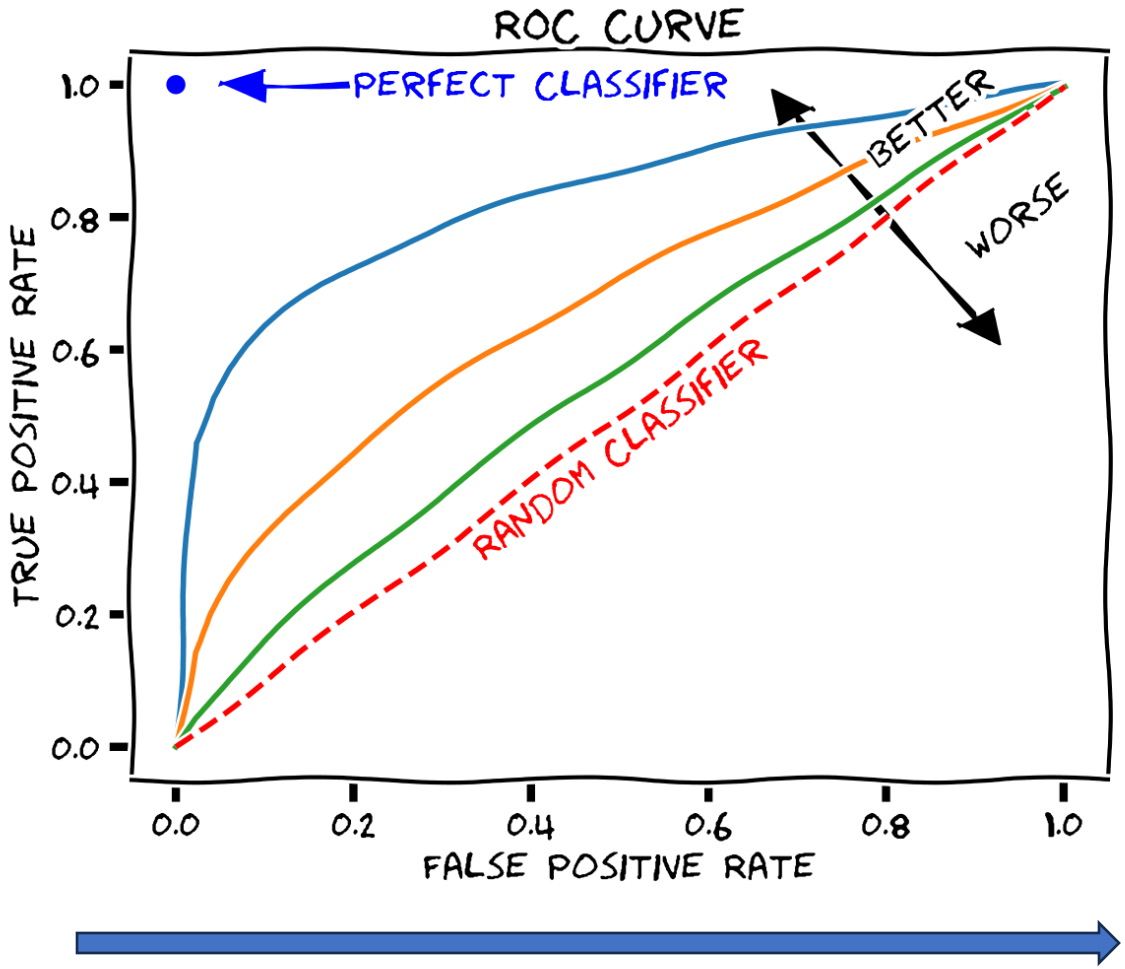
(b) Domain indicators, avg. overlap 4,0%

Even when tracking the same threat actors, two paid CTI feeds have minimal overlap of indicators, suggesting very low overall coverage of maliciousness.

https://www.usenix.org/conference/usenixsecurity20/presentation/bouwman

Cost of compromise vs disruption

Attackers Disrupted more often



Employees disrupted more often

Detecting Malware
(later maliciousness)

1. **Signature-based (what it is)**
   --> file hash, patterns of bytes or similar
2. **Static + dynamic analysis (what does it look like/do when it runs)**
   --> analyse source code or run it in a VM and see what it does
3. **Behavioural detection (what it does on the real endpoint)**
   -->EDR model: monitor live activity on actual machines and flag suspicious sequences of actions
   No longer just about malware.

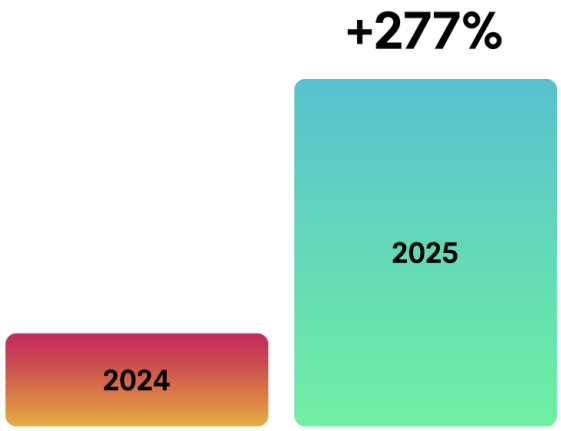
EDR Detection Signals

- Monitor executables, especially if file just downloaded from internet
  - Is cmd.exe less likely to be a legit employee.... unless they're in tech
- Protocols/services rarely used by "normal" employees are also suspicious
- RMM tools just downloaded should be monitored
  - Especially if not the corporate solution
- Mass (any?) encryption
- ...

Top 5 Most Frequently Seen MITRE ATT&CK Sub-Techniques

Rank	Technique	Percent
1	T1059.001: PowerShell	26.2%
2	T1021.002: SMB/Windows Admin Share	23.3%
3	T1021.001: Remote Desktop Protocol	22.6%
4	T1070.004: File Deletion	21.7%
5	T1569.002: Service Execution	19.0%

https://services.google.com/fh/files/misc/m-trends-2025-en.pdf



Attackers are ditching traditional hacking tools and abusing trusted software that blends into normal business workflows—RMM abuse alone is up 277% year over year.

https://www.huntress.com/resources/2026-cyber-threat-report

EDR Detection Signals

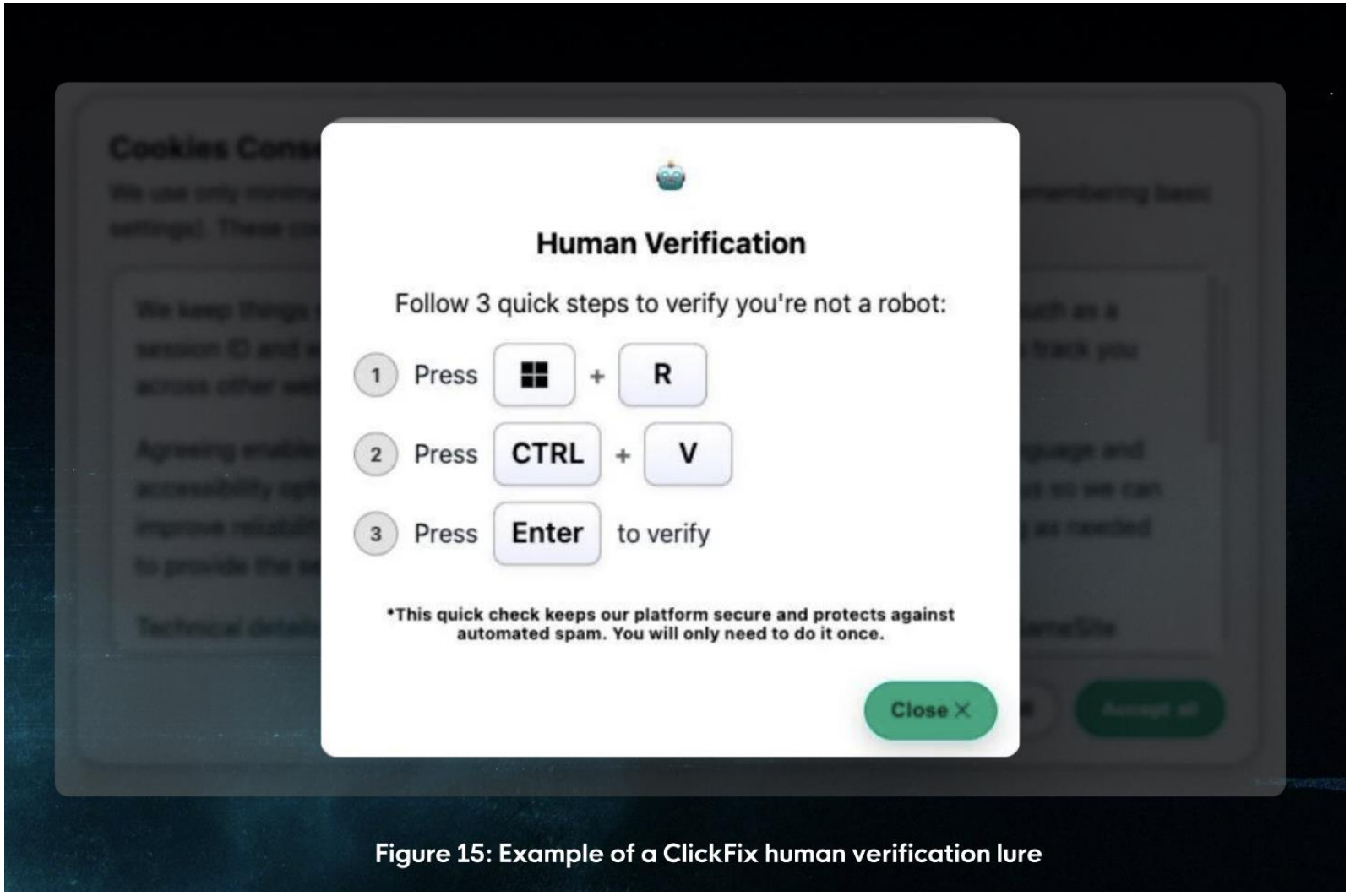


Figure 15: Example of a ClickFix human verification lure

https://www.huntress.com/resources/2026-cyber-threat-report

The suite of detection tools

In a typical network

- Anti-virus/end-point
- Email server
  - Spam, malware, fraud etc
- Identity
  - SSO, email, VPN etc
  - New + emerging
- Firewall/network
  - Data loss prevention

In niche settings

- Credit card fraud
  - ATM
  - Terminal
  - Web payments
- LLM
  - Distillation, malicious use
- Etc etc etc

Detection is distributed across products/activity/logs.

Approaches to Detection

- Correlating/chaining signals together
  - Only fire if X then Y then Z
- Statistical approaches
  - Machine learning
  - Anomaly detection
- Escalating to a human to investigate further
  - Security operations centres
  - Digital forensics and incident response

If all else fails

- Attackers may alert you
  - Ransom note
- Business partners may alert you
  - If your BEC led to their fraud
  - If your compromise led to theirs
- Third-parties may alert you
  - Cybercrime researchers
    - Dark web data
    - Infostealers
  - Credit card provider
  - ISP and intel community
  - Your insurer

This threat intel has v low false positive rate



Early Warning

Free malicious activity notifications from the NCSC for UK organisations.



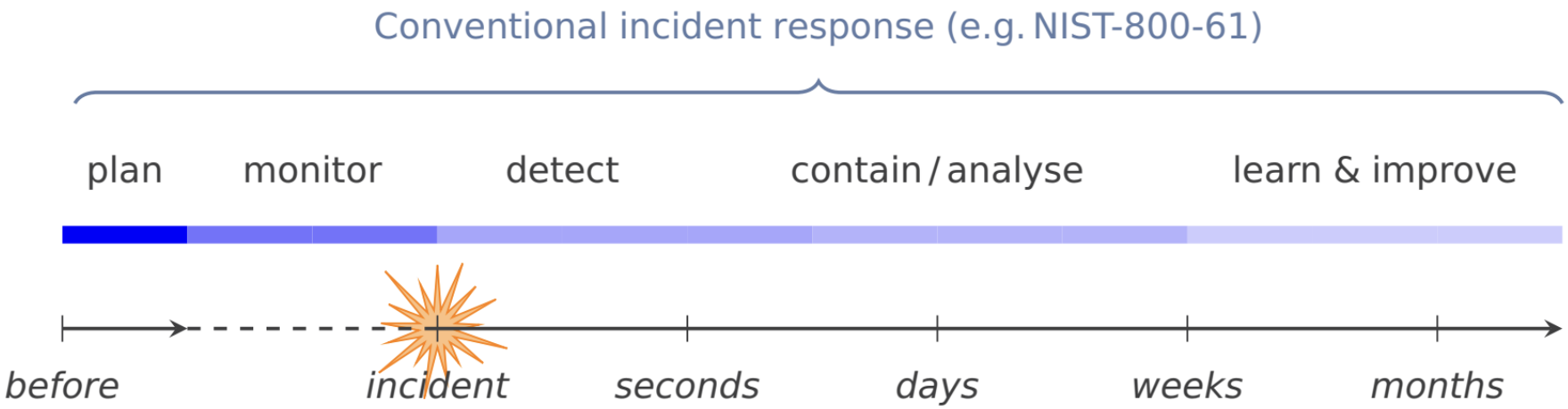
Detection and Response: Stopping suspicious activity

Automated Actions by Tools

- Notify
  - Flash a warning to the user
  - Send an alert to the SOC
- Evict/contain
  - Identity: invalidate a login token or force a credential reset
  - Network: end a network connection or block an address
  - Machine: block an executable, end a process etc

But what if this goes wrong...

Traditional Incident Response



- Incidents as a fact of life
- Detection, response, analysis, forensics, learning as one process

… what goes wrong?

What goes wrong with NIST-800-61

- Many organizations can't afford detection tools, let alone dedicated IT staff
  - Detection = crisis
- Reputation + legal management
  - Better not to publish findings
  - May damage reputation or litigation
- Need to protect findings with legal privilege
  - In the US, UK and beyond



"One of the most important parts of incident response [...]: learning and improving."

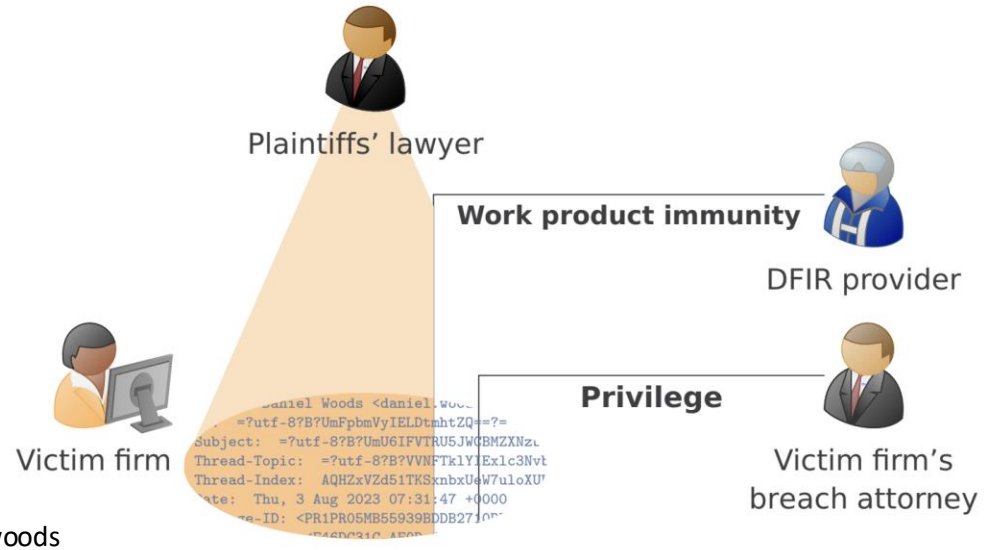
NIST 800-61, p. 38

"Insurers accumulate data [...] which they mine to improve risk assessment and suggest best practice mitigation strategies to their clients."

Anderson et al. 2008 (ENISA), p. 40

"The IT industry does not have strong processes for extracting lessons learned and publishing them when incidents occur."

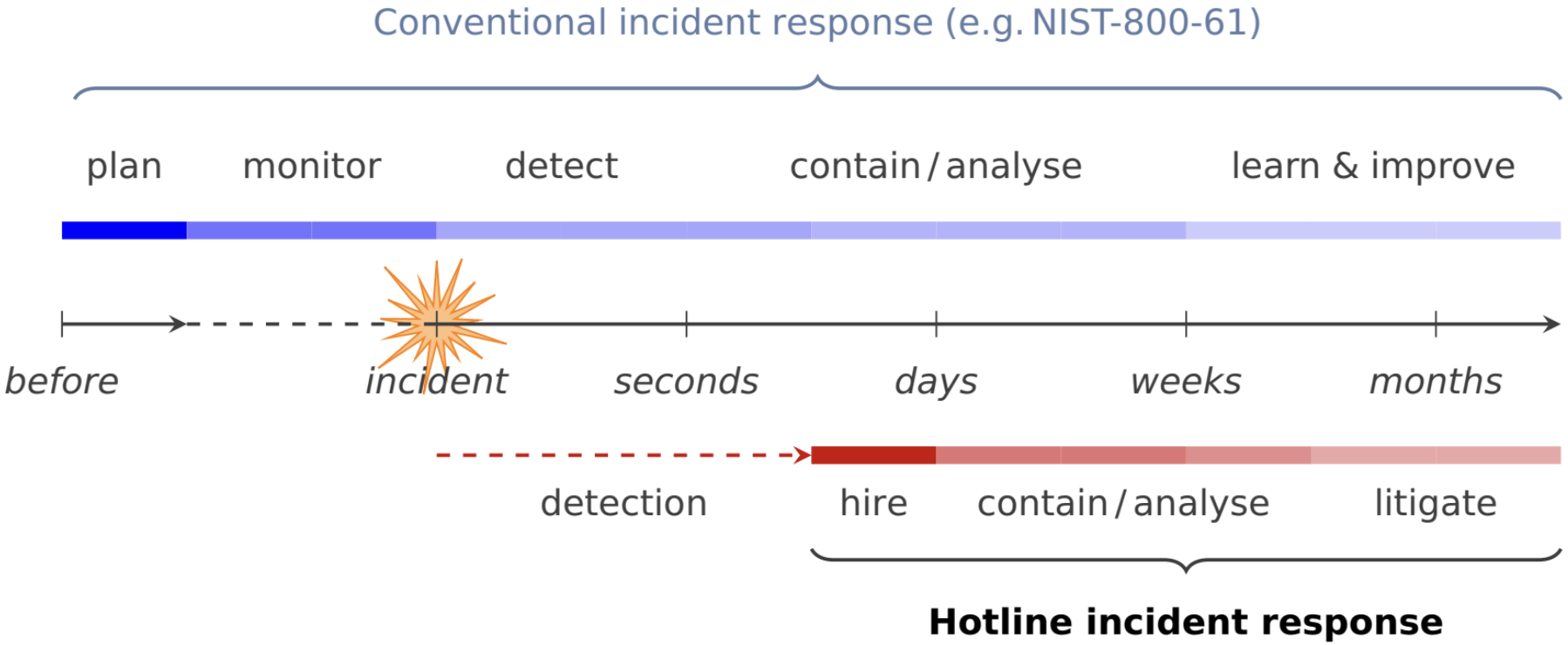
Knake, Shostack & Wheeler 2021 (Harvard Belfer Center)



https://www.usenix.org/conference/usenixsecurity23/presentation/woods

Hotline Incident Response

- Incident response as crisis vs as a process



Woods, D. W. and Böhme, R. Incident Response as a Lawyers' Service. *IEEE Security & Privacy*, **20**, 2 (2022), 68–74.

Digital Forensics

- Field grows out of law enforcement
  - Own self-contained field
  - Hollywood aspect of InfoSec
- Goals
  - Contain the threat actor
  - Chain of custody
    - Image machines/network
  - Identify what data was accessed
  - …
- Only possible with good logs
  - Often missing due to economics and TAs
  - Gaps across products

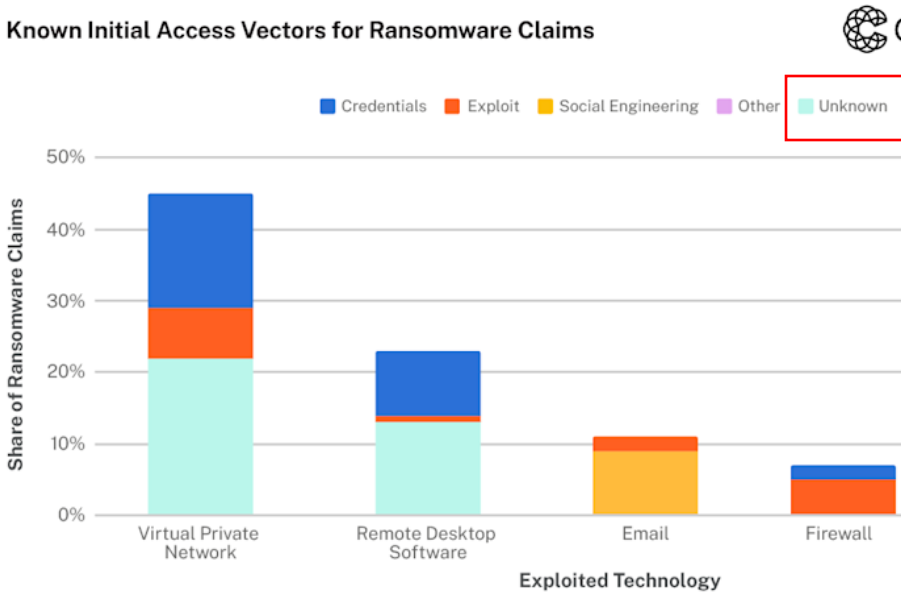
Microsoft offers free security logs amid backlash from State Department hack

Federal officials and rivals blasted the company for charging customers for additional security features.

Published July 19, 2023

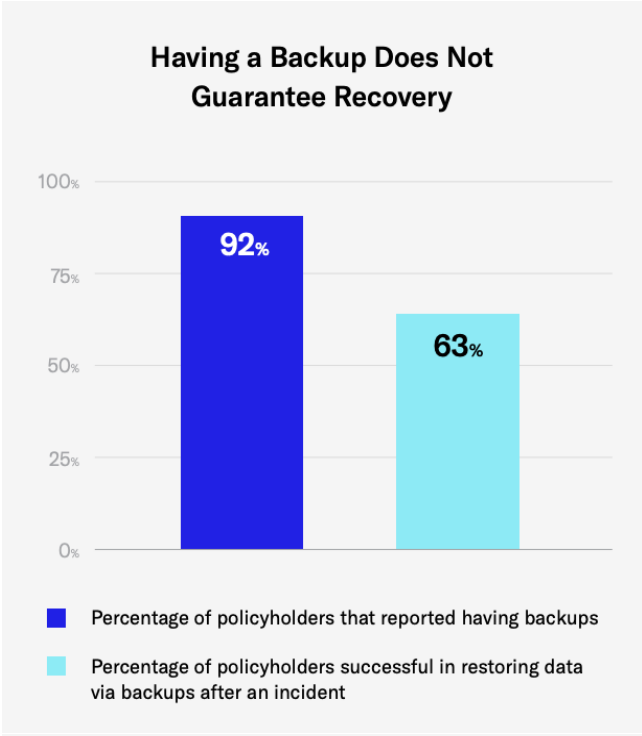


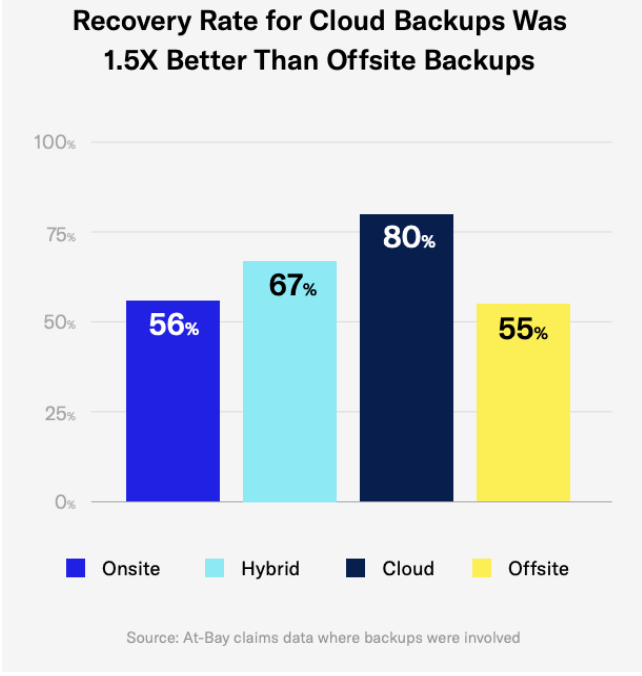
David Jones
Reporter



Incident Response

- About getting systems back online as fast as possible
  - In tension with investigative goals?
- Rebuilding systems
  - Easy to buy a new machine
  - Harder to rebuild a complex network
    - Maersk & Not Petya, power cut at domain controller in Ghana saved the day
- Restoring data relies on backups
  - Often corrupted or too slow
  - Targeted by attackers

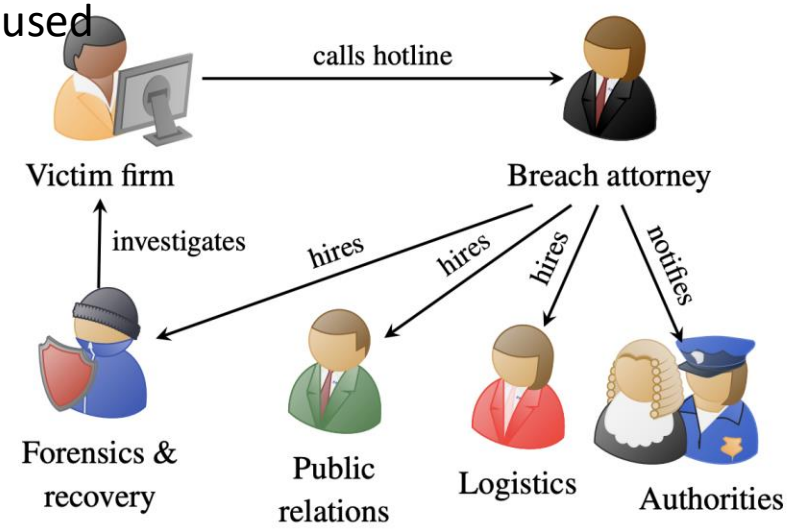




Source: https://www.at-bay.com/wp-content/uploads/2023/10/At-Bay_Backup-Breakdown.pdf

Non-Technical Responders

- Need to tell your insurer, your regulator, your...
- Lawyers
  - To advise on breach notification requirements
  - To make sure no-one says anything that'll be used against them
- E-discovery
  - To mine stolen data and identify PII
  - Can be very very expensive (at least pre-AI)
- Logistics providers
  - Sending breach notices by mail
  - Spinning up comms infra
- Public relations
  - To communicate with the public/customers
- And then...
  - Psychologists?
  - Temporary relocation?



Summary

- Leaky perimeters: do we block, detect, or respond?
- Blocking is most effective at preventing compromise
  - But also adds the most friction
  - Allow-listing most effective and most friction
- Instead of blocking, we hope detection and response catches maliciousness
  - Reliably detect the obvious stuff, but then attackers start to act like defenders (e.g. RMM tools)
  - We still impose user friction by blocking or alert fatigue
  - Now we might also be too slow to stop the breach
- Incident response is messy.
  - Used to be an extension of detection, now new people come in
  - Hard to undo bad preparation (e.g. no logs, no backups)
  - IR has non-technical components that are arguably as important