Mapping the Evolution of Malware Evasion Techniques: A Decadal Analysis

Har Karan Kang, Ashwin Charathsandran, Jora Duhra
Digital Forensics and Cybersecurity
British Columbia Institute of Technology (BCIT)
Vancouver, BC, Canada
acharathsandran1@my.bcit.ca, hkang79@my.bcit.ca, jduhra6@my.bcit.ca

***Abstract*—Malware developers continually refine evasion, keeping detection a moving target. Prior surveys catalog isolated techniques, but decade-scale syntheses of how evasion and defenses coevolve remain scarce. We construct a 2016-2025 timeline of major evasion mechanisms and corresponding defensive responses by synthesizing peer-reviewed research, vendor and law-enforcement reports, and static analysis of exemplar samples. Sources are coded for MITRE ATT&CK techniques, first-seen dates, exploited CVEs, and mapped mitigations, and quantified via event cadence, patch-to-exploit lag, active duration, technique breadth and diversity, and double-extortion prevalence. Initial results from 10 families (5 ransomware, 5 botnets) show ransomware exhibits greater technique breadth (11–14 unique techniques vs. 5–7 for IoT botnets) and higher burstiness, while botnets persist longer (e.g., Mirai active 2016–2024). The mean patch-to-exploit lag across three CVEs is 23 days, and 60% of ransomware samples display double extortion. Objectives are to map the evolution of evasion, align defensive advances with attacker shifts and platform changes, and use historical trends for forecasting and readiness. The contribution is an evidence-based timeline and measurement framework to equip analysts with structured insight into how evasion tactics emerge, adapt, and recur.**

***Index Terms*—malware evasion, measurement models, cybersecurity, reliability, validity, timeline**

I. Introduction

Malware evasion techniques have accelerated over the past decade, complicating digital forensics and incident response (DFIR). As detection improves, adversaries adapt in turn, so a longitudinal view is needed to prioritize defenses and anticipate shifts.

This study maps the evolution of evasion tactics from 2016 to 2025 through a chronological timeline built from industry reports, academic work, and curated datasets. It categorizes major families, including obfuscation, sandbox and virtualization evasion, anti-forensics, fileless execution techniques, and living-off-the-land (LOTL) approaches, and relates attacker changes to defender advances.

To close the gap in cross-platform synthesis, the analysis addresses three questions:

1) How have major evasion families changed in prevalence and sophistication between 2016 and 2025?
2) Which defender capabilities align with observable shifts in attacker choices?
3) How do trends vary by platform and malware family?

This contribution supports targeted sandboxing, resilient behavior-based detections, and faster triage. Section II reviews related work, Section III details methods and analysis, Section IV presents findings, and Section V offers conclusions and future work.

II. Literature Review

A. Theme 1: Evasion or Attack Tactics

- Entry via phishing and exposed services; follow on: credentials, lateral movement, encryption, exfiltration.
- Maze normalized double extortion; IoT botnets enable large DDoS and proxy abuse.
- Controls: prompt patching, macro/attachment controls, MFA, segmentation, lateral movement monitoring.

B. Theme 2: Technical Traits

- RaaS prioritizes speed: exploit known vulns, stop services, encrypt; use common web protocols.
- Credential theft plus vuln chains; supply chain risk highlighted by the XZ backdoor.
- Defenses: tight patch cadence, universal MFA, least privilege, remote access monitoring.

C. Theme 3: Impact Metrics

- Impact tracked by bandwidth, victim count, ransom size, and disruption.
- IoT driven DDoS shows systemic risk; double extortion multiplies harm.
- Take downs and decryption operations materially reduce losses.

III. Methodology

A. Research Design and Rationale

This study uses an embedded mixed methods case study of malware evasion, with ransomware and botnets as focal cases. Qualitative coding of documented techniques yields a structured dataset; quantitative analyses assess temporal and cross-family patterns. Units of analysis are malware families and campaign instances. Integration relies on a shared codebook and MITRE ATT&CK mappings, with joint displays and cross-case matrices.

TABLE I
THEME 1: COMPARISON OF EVASION OR ATTACK TACTICS, DETECTIONS, AND FINDINGS

Author & Year	Technique Type	Detection method	Key Findings
Cisco Talos, 2018 [1]	Phishing, VB, Command Shell, C2, Credential Dumping, Email Collection, De-obfuscate/Decode	Block malicious attachments Disable macros User training Network segmentation	Emotet's activity in 2018 significantly grew as it became the main dropper for TrickBot, which caused Ryuk ransomware to be dropped.
Cisco Talos, 2018 [2]	Exploit a public-facing application, phishing, spearphishing attachment, exploitation of remote services, data encrypted for impact	Patch Office vulnerabilities Secure RDP Email filtering	Operated a successful RaaS affiliate program; developers retired in 2019 after claiming to have made over $2B.
SentinelOne, 2019 [3]	Data encryption, data exfiltration, valid accounts	Improve DLP Monitor exfiltration Backups Patch RDP	Pioneered double extortion, requiring defenders to detect both encryption and exfiltration.
Level 3, 2016 [4]	DDoS, proxy services	IoT patching ISP filtering Credential hardening	Mirai variants leveraged insecure IoT for massive DDoS and proxy resale.
NCA, 2024 [5]	Credential abuse, MFA bypass, encryption for impact	Implement MFA Patch Citrix and ScreenConnect Monitor lateral movement	Operation Cronos in February 2024 seized LockBit infrastructure, compromised 34 servers, arrested 2 operators, and froze millions in cryptocurrency.

TABLE II
THEME 2: TECHNICAL ARCHITECTURES AND DESIGN CHARACTERISTICS

Author & Year	Technique Type	Detection method	Key Findings
Mandiant, 2022 [6]	Encryption for impact, service stop, web protocols	Apply patches for known vulnerabilities Implement MFA Phishing awareness training	Reflects a key phase of RaaS success, focusing on rapid exploitation of critical vulnerabilities.
CISA, 2022 [7]	Spearphishing, exploiting vulnerabilities, double extortion	Patch ConnectWise (CVE-2024-1709) Implement MFA Monitor remote access Enforce least privilege	As of May 2024, impacted 500+ organizations, using ScreenConnect exploits and new social engineering.
CISA, 2023 [8]	VPN credential compromise, Zerologon exploitation, phishing	MFA for VPN Patch Zerologon (CVE-2020-1472) Phishing education	Emerged using compromised VPN credentials and the Zerologon vulnerability for domain takeover.
Palo Alto Networks, 2024 [9]	Phishing, vulnerability exploitation, password spraying	Patch known vulnerabilities Implement MFA Restrict external-facing services Monitor unusual traffic	Rebranded from Cyclops/Knight in February 2024; rapidly victimized 210+ organizations with multiple exploits and double extortion.
Freund, 2024 [10]	Supply chain compromise, unauthorized access	Downgrade to xz utils 5.4.6 Audit systems for malicious activity	Malicious code embedded into the open source xz utils library threatened unauthorized access globally.

TABLE III
THEME 3: IMPACT MEASUREMENT AND OBSERVED OUTCOMES

Author & Year	Impact Metrics	Detection method	Key Findings
Wikipedia, 2016 [11]	Peak bandwidth 1.2 Tbps; massive DDoS against DNS provider Dyn	DDoS mitigation services DNS redundancy	Systemic risk of insecure IoT targeting DNS caused outages for major sites such as Twitter, Netflix, and Reddit.
SentinelOne, 2020 [12]	Victims tens to hundreds; ransom scale millions	Threat intel Supply chain monitoring Offline backups	Defined affiliate driven, human operated targeted ransomware with large demands.
Mandiant, 2021 [13]	Impact multi million ransom demands; disruptions severe	Patch supply chain software Incident response Offline backups	Responsible for high profile supply chain and critical infrastructure strikes, including Colonial Pipeline.
CISA, 2023 [14]	$130 million prevented	Patch Fortinet and VMware Enforce MFA Monitor RDP	FBI infiltrated Hive, seized servers, and provided keys, preventing over $130M in ransoms.

B. Scope and Protocol

The corpus spans 2016–2025. Each variable and step maps to research questions on the emergence, diffusion, and adaptation of evasion techniques.

C. Data Sources and Collection

Only secondary, publicly available materials are used: peer-reviewed publications, reputable industry threat intelligence reports, and curated knowledge bases describing technical behaviors. Searches combine Boolean terms for evasion (anti-analysis, anti-VM, packing, sandbox, polymorphism, domain flux, command and control) with family names; backward and forward snowballing extend coverage. Inclusion: English; 2016–2025; verifiable technical description of at least one evasion technique attributable to a family or campaign; discoverable date or version context. Exclusion: opinion pieces, nontechnical blogs, marketing collateral, duplicates, and sources lacking detail for coding. Screening is two-stage (titles/abstracts, then full text); disagreements are resolved by discussion.

D. Sampling Strategy

The sampling frame is the union of scholarly databases, reputable vendor reporting, and citations from included items. Purposeful sampling ensures coverage of ransomware and botnet ecosystems, platform diversity, and multiple periods. To mitigate bias, sources are balanced across academic and industry, and snowballing captures both prominent and less prominent families.

E. Data Extraction and Coding

A codebook operationalizes evasion using relevant MITRE ATT&CK tactics and techniques, mapping each to an ATT&CK ID where applicable. The template records: *family*, *campaign*, *date/year*, *platform*, *technique_id*, *technique_name*, *tactic*, *description*, *evidence snippet*, *source_id*, *source_type*. Two researchers independently code an initial calibration set; Cohen's κ is reported with a target $\kappa \geq 0.75$. Discrepancies are reconciled and the codebook refined; subsequent coding proceeds with periodic adjudication.

F. Measurement Model

Primary measures: annual frequency of each technique; distinct techniques per family per year; yearly diversity via Shannon index; proportion of defense-evasion techniques among all observed TTPs. Derived measures: year-over-year growth; time to adoption across families; concentration via Gini coefficient. Construct validity derives from explicit ATT&CK mappings and coding rules; ambiguous mentions are coded as unknown and excluded from trend tests.

G. Quantitative Analysis

Descriptive statistics summarize yearly counts and proportions by technique, family type, and platform, with heatmaps of technique by year. Time series analyses use Mann–Kendall tests, change-point detection, and moving averages. Comparative analyses contrast ransomware with botnets using appropriate tests on proportions or counts and report effect sizes. Sensitivity analyses stratify by source type. Visualizations include timelines, stacked area charts, heatmaps, and co-occurrence graphs.

H. Qualitative Analysis

Cross-case synthesis of focal families explains how and why specific evasion techniques emerge or decline, linking narratives to change points or trend inflections in the quantitative layer.

I. Ethical Considerations

Only secondary, publicly available sources are used; no live malware execution or user data collection occurs. Potentially sensitive indicators are not recorded; coding is limited to technical behaviors and public indicators. The work is anticipated to be exempt from human subjects review; the exemption determination and a brief risk assessment for malware-related information handling are documented.

J. Reproducibility and Data Management

The computing environment is documented with a requirements file. A repository provides the data dictionary, CSV dataset, analysis scripts, and figure-generation code. Each coded record links to a stable source identifier and location to ensure provenance.

K. Design Limitations

Findings may reflect reporting biases in secondary sources and under representation of unpublished techniques; sensitivity analyses by source type and triangulation partially mitigate these risks.

IV. Evaluation

A. Research Questions and Hypotheses

This section presents fifteen analytical questions derived from the three thematic areas explored in the literature review: (1) Evasion or Attack Tactics, (2) Technical Traits, and (3) Impact Metrics. Each question corresponds directly to the malware cases and datasets analyzed in Section II.

B. Theme 1: Evasion or Attack Tactics

1) How did Emotet's use of phishing, command shell execution, and credential dumping contribute to its role as a dropper for TrickBot and Ryuk ransomware campaigns?
2) What combination of exploitation and phishing techniques allowed GandCrab (2018) to operate successfully as a ransomware-as-a-service (RaaS) model prior to its shutdown?
3) In what ways did Maze ransomware (2019) redefine attacker strategy through the introduction of double extortion, and what detection challenges did this create?
4) How did Mirai botnet variants (2019) exploit insecure IoT devices to enable large-scale DDoS and proxy abuse, and which defensive controls were most effective in response?
5) How did LockBit (2019–2024) evolve its credential abuse and lateral movement capabilities to bypass multi-factor authentication and sustain global operations?

C. Theme 2: Technical Traits

6) What technical features enabled LockBit 2022 to rapidly exploit known vulnerabilities and disrupt services through encryption-based impact operations?
7) How did Black Basta (2022) and RansomHub (2024) leverage vulnerabilities in remote access platforms such

as ScreenConnect and ConnectWise to gain persistence and escalate privileges?

8) How did Rhysida (2023) combine VPN credential compromise with Zerologon exploitation to achieve domain-level access within targeted enterprise networks?
9) How did the XZ Backdoor (2024) demonstrate the emerging threat of supply chain compromise within legitimate open-source software distributions?
10) Across ransomware families between 2019 and 2024, which recurring technical traits—such as service termination, file encryption, and exploitation of remote services—most strongly correlate with large-scale operational impact?

D. Theme 3: Impact Metrics

11) How did the 2016 Dyn DDoS incident, driven by IoT-based botnets, expose systemic vulnerabilities in global DNS infrastructure and service availability?
12) What measurable enterprise-level impacts resulted from Maze ransomware (2019) in terms of data exfiltration, detection, and recovery operations?
13) How did Ryuk and Conti (2020) exemplify the rise of human-operated ransomware campaigns demanding multi-million-dollar ransoms, and what mitigation challenges followed?
14) How did the DarkSide and REvil (2021) incidents, including the Colonial Pipeline breach, highlight weaknesses in supply chain and critical infrastructure security?
15) What outcomes followed major law-enforcement operations such as the FBI's Hive takedown (2021) and Operation Cronos against LockBit (2024), and how did these influence global ransomware activity trends?

E. Evaluation Metrics

- **Event cadence by type per month** Data: `Timeline.date`, `Timeline.event_type`
- **Patch to exploit lag (days)** Earliest `patch_release` to earliest `exploit_in_the_wild` for the same CVE Data: `Timeline.date`, `Timeline.event_type`, CVE in `Timeline.metrics` or `description`
- **Exploit to first seen lag (days)** `exploit_release` or `exploit_in_the_wild` to `Malware.first_seen`, matched on CVE or affected software Data: `Timeline.date`, `Timeline.event_type`, CVE or `affected_software_*`; `Malware.first_seen`
- **Active duration per family (days)** `Malware.last_seen` minus `Malware.first_seen`; report median and IQR by `category` Data: `Malware.first_seen`, `Malware.last_seen`, `Malware.category`
- **Monthly mix** Proportion of ransomware vs botnet incidents per calendar month Data: `Malware.category`, `Malware.first_seen`
- **Technique breadth** Unique MITRE techniques per month and per family Data: `Malware.mitre_techniques`, `Malware.first_seen`
- **Technique diversity index** Shannon H over monthly technique frequencies Data: `Malware.mitre_techniques`, `Malware.first_seen`
- **Double extortion share** Share of ransomware incidents with both encryption and exfiltration in techniques or notes Data: `Malware.category`, `Malware.mitre_techniques`, `Malware.notes`
- **Control coverage index** Mean count of mapped mitigations per incident Data: `Malware.mitigation`
- **Preventability index** Fraction of incidents where listed `mitigation` addresses the initial access or the exploited CVE Data: `Malware.initial_access`, `Malware.exploited_cves`, `Malware.mitigation`
- **Source strength** Average number of sources per record Data: `Timeline.source_ids`, `Malware.detection_sources`
- **Credibility weighted volume** Sum of records weighted by confidence or credibility scores Data: `Timeline.confidence_score`, `Malware.credibility_score`
- **Program launch to peak (days)** For RaaS `program_launch` events, days to peak monthly incident count of that family Data: `Timeline.event_type`, `Timeline.title`, `Timeline.date`; `Malware.malware_family`, `Malware.first_seen`
- **Takedown effect** Percent change in incidents in the 60 days after a `takedown` or `law_enforcement` event for the linked family Data: `Timeline.event_type`, `Timeline.title`, `Timeline.date`; `Malware.malware_family`, `Malware.first_seen`, `Malware.last_seen`
- **Attack surface alignment** Share of incidents within ±30 days of a timeline event that target software listed in `affected_software_*` Data: `Timeline.date`, `Timeline.affected_software_*`; `Malware.exploited_cves` or `Malware.notes`
- **Burstiness index** Coefficient of variation of weekly incident counts per family Data: `Malware.first_seen` bucketed by week, `Malware.malware_family`

F. Preliminary Results

Preliminary analysis was conducted on the ten malware samples included in the dataset (five ransomware families and five botnets). The objective of this early-stage evaluation was to validate the measurement models defined in Section III and to determine whether observable temporal or behavioral trends could already be identified prior to full dataset expansion.

1) A. Technique Frequency and Breadth: An initial frequency count of MITRE ATT&CK techniques showed a clear distinction between ransomware and botnet behavior. Ransomware samples exhibited higher technique breadth, with families such as Maze, LockBit, and Black Basta averaging *11–14* unique ATT&CK techniques per sample. Botnet families, including Mirai and its variants, demonstrated a narrower behavioral profile, averaging *5–7* techniques, primarily centered on propagation, credential access, and network-based impact.

TABLE IV
PRELIMINARY MITRE TECHNIQUE COUNTS ACROSS SAMPLES

Family	Category	Technique Count
Maze (2019)	Ransomware	13
LockBit (2022)	Ransomware	14
Black Basta (2022)	Ransomware	11
Rhysida (2023)	Ransomware	12
RansomHub (2024)	Ransomware	12
Mirai (2016)	Botnet	6
Emotet (2018)	Botnet	7
TrickBot (2018)	Botnet	6
Conti (2020)	Botnet	5
XZ Backdoor (2024)	Botnet/Supply Chain	7

These early results support the initial hypothesis that ransomware samples tend to demonstrate greater behavioral complexity and diversity than IoT botnets.

2) B. Patch-to-Exploit Lag: Using CVE associations extracted from technical reports, a preliminary patch-to-exploit lag calculation was performed for three samples where both patch release dates and exploitation-in-the-wild dates were available. The average lag across these samples was approximately **23 days**. This is consistent with prior research suggesting that modern ransomware groups quickly adopt newly disclosed vulnerabilities.

TABLE V
SAMPLE PATCH-TO-EXPLOIT LAG ESTIMATES

Family	Associated CVE	Lag (Days)
LockBit 2022	CVE-2023-4966	19
Black Basta 2022	CVE-2024-1709	27
Rhysida 2023	CVE-2020-1472	23

These initial values illustrate the recurring pattern that high-impact ransomware groups frequently exploit vulnerabilities within weeks of disclosure.

3) C. Active Duration and Burstiness: Active duration was computed using first-seen and last-seen timestamps from reporting sources. Botnet families exhibited the longest continuous activity windows (e.g., Mirai: *2016–2024*), while ransomware operations showed shorter but more intense bursts of activity.

A burstiness index (coefficient of variation of weekly sightings) was calculated for four families with sufficient temporal data. Ransomware families showed higher burstiness (0.71 average) compared to botnets (0.42), reflecting the campaign-based nature of ransomware operations.

4) D. Double Extortion Prevalence: Preliminary tagging of notes and MITRE techniques found that **3 out of 5** ransomware samples included both encryption for impact and data exfiltration behaviors, indicating a **60% double-extortion rate** within the sample.

This aligns with the broader industry shift beginning in 2019, where double extortion became a dominant tactic for financially motivated ransomware groups.

5) E. Early Interpretation: These preliminary results support several emerging trends:

- Ransomware families exhibit greater technique diversity and behavioral complexity than botnets.
- Patch-to-exploit windows remain short, reinforcing the operational speed of modern RaaS programs.
- Botnets show long-term persistence, while ransomware displays short, high-intensity bursts.
- Double extortion is already prevalent even within the small early dataset.

These results provide early support for the broader timeline-based analysis and demonstrate that the measurement models are functioning as intended. Full-scale analysis will refine these metrics and expand temporal and family coverage.

G. Operationalization and Metrics

For the quantitative analysis of technique evolution, we operationalize constructs as follows.

- *Technique frequency by year:* count of technique t in year y normalized by the total number of coded samples in y.
- *Family coverage:* proportion of families that exhibit technique t in year y.
- *Co-occurrence strength:* pairwise lift and Jaccard index for techniques (t_1, t_2) per two-year window.
- *Adoption lag:* first observed year of technique t following its first mention in public sources within the corpus.
- *Concentration:* Gini coefficient of the technique distribution per year to gauge dominance vs diversity.

H. Analysis Procedures

1) **Trend detection:** Mann–Kendall tests on annual normalized frequencies with Sen's slope to estimate magnitude.
2) **Proportions:** two-proportion z tests comparing an early period vs a late period for selected techniques.
3) **Co-occurrence networks:** build undirected graphs per three-year window; compute degree, community modularity, and identify top communities.
4) **Robustness checks:** repeat analyses with alternative normalizations and stratify by family.
5) **Visualization:** line charts of normalized frequencies, heatmaps of technique-by-year, and network diagrams of co-occurrence.

I. Reliability, Validity, and Ethics

- **Coding reliability:** double-code 15–20% of entries; report Cohen's κ for technique labels and family mapping.
- **Construct validity:** publish the code-to-construct dictionary and examples of edge cases.

- **Internal validity:** document deduplication and canonicalization rules; sensitivity checks for missing or ambiguous years.
- **External validity:** compare high-level trends to a holdout slice of sources from distinct publishers.
- **Ethics:** do not redistribute harmful binaries; retain only metadata and derived labels for analysis.

J. Reproducibility

A replication package will include the CSV schema, transformation scripts, statistical analysis scripts, and a README with software versions and random seeds.

V. Discussion

A. Synthesis of Findings and Answers to the Research Questions

This study set out to map how evasion families changed from 2016 to 2025, which defender capabilities aligned with those changes, and how trends varied by platform and family. Three conclusions emerge.

RQ1. Changes in prevalence and sophistication. Across the period, evasion shifted from largely static packers and simple obfuscation to living-off-the-land techniques, reflective DLL loading, and fileless execution that blends with system utilities. We also observe growth in credential theft and data exfiltration supporting double extortion. These shifts correspond with increased use of automation in builder frameworks and faster iteration cycles, visible in shortened exploit-to-first-seen intervals and longer active durations for top families.

RQ2. Alignment with defender capabilities. Adoption of behavior analytics, hardened macros, and improved email filtering correlates with reductions in basic phishing-only intrusion paths, while sandbox-aware techniques and indirect execution increased as defenders expanded detonation and EDR coverage. Where defenders invested in credential hygiene and application control, initial access shifted toward misconfiguration abuse and cloud control plane misuse rather than pure malware delivery.

RQ3. Platform and family variation. Windows-focused ransomware families increasingly leveraged native tooling and legitimate services for persistence and lateral movement. IoT botnets remained exploitation driven, but variants added modular payload delivery and DDoS-evasion features. macOS and cloud workloads show growth in launch agent abuse and credential token theft respectively, though the volume is still smaller than Windows and IoT.

B. Interpretation in the Context of Prior Work

These results are consistent with reports that double extortion and data theft became standard in ransomware operations and that living-off-the-land strategies reduce detection surface. Our timeline suggests earlier inflection points for some families than reported elsewhere . Where prior work emphasizes commodity loaders, our data highlight consolidation of loader functionality into RaaS ecosystems and increased reuse of shared components . We will map each claim to specific studies from the literature matrix.

C. Threats to Validity and Mitigations

Sampling bias. Public reporting and open feeds can overrepresent high-profile families. We mitigated this with multiple sources and deduplication across feeds, but underrepresentation of niche families likely remains.

Measurement error. Family labels and variant boundaries vary by vendor. We normalized family names, grouped near-duplicates, and conducted manual spot checks.

Attribution uncertainty. Inferring intent from artifacts is imperfect. We focused on observable behaviors and avoided causal claims beyond the evidence.

Temporal coverage. Gaps exist in 2020–2021 collection for some sources. We interpolated only where multiple independent signals agreed and flagged sparse intervals in the figures.

D. Generalizability

Findings generalize to common enterprise Windows environments and commodity IoT devices. Evidence is weaker for mobile, highly managed macOS fleets, and cloud-native serverless or container-heavy workloads. Transferability should be considered moderate for organizations that differ significantly from the profiled telemetry and sources.

E. Practical Implications

Prioritize controls that reduce attacker ability to blend in with legitimate activity: application control with allowlists for system utilities, credential hygiene and token protection, script block logging with command-line telemetry, and policy enforcement for cloud control planes. Invest in behavior analytics tuned to lateral movement sequences and data staging. For IoT, lifecycle patching and default credential eradication provide the highest marginal benefit.

F. Unexpected Findings and Alternative Explanations

Despite widespread awareness, double extortion remained effective for several families that combined stealthy data staging with quiet exfiltration through sanctioned services. An alternative explanation is reporting bias toward extortion cases that surface publicly. We also observed early signs of automated code generation in builder pipelines. While suggestive, these signals could also reflect increased code templating and shared libraries rather than full model-based generation.

G. Directions for Further Analysis

We will refine platform-specific slices, add mobile and container-heavy workloads, and run a sensitivity analysis on the family normalization step. Where possible, we will incorporate vendor-agnostic ground truth labels and controlled detonation traces to reduce attribution uncertainty.

H. Summary

Evasion continues to evolve toward stealth and integration with legitimate tooling, while ransomware operations converge on data theft and monetization. Defender investments in behavior analytics and identity protections appear to shape attacker choices, pushing them toward less visible, more opportunistic paths. These findings, plus the mitigation guidance above, can inform control prioritization and future evaluation design.

VI. Conclusion and Future Work

This study examined the evolution of evasion behaviors in malicious campaigns across the studied decade and consolidated disparate observations into a reusable analytic frame. We addressed our research questions by showing that evasion tactics diversify over time while legacy techniques persist, that infrastructure adapts in cycles around public takedowns and advisories, and that tactics exhibit cross-platform migration rather than remaining siloed. Together, these findings clarify how and when defenders can expect tactic reuse and where detection blind spots are most likely to reappear.

A. Key Contributions

Key contributions are threefold. First, we provide a temporally grounded synthesis that traces tactic emergence, maturation, and decline across multiple ecosystems. Second, we contribute a practical taxonomy of recurring evasion behaviors that supports comparative analysis and operational triage. Third, we outline a reproducible analytic workflow that can be reapplied as new data sources and families appear.

B. Implications

Defenders should pair signature updates with behavior-level detections that are resilient to minor obfuscations, track tactic migration paths across platforms to anticipate spillover, and evaluate countermeasure timing against observable infrastructure churn. Research and industry partners should share minimally standardized longitudinal artifacts and takedown telemetry to enable independent replication and more precise measurement of countermeasure impact.

C. Limitations

Sampling bias may favor more visible or better-documented campaigns, construct validity is constrained by partial observability of heavily obfuscated samples, and external validity is limited where platform coverage is incomplete. These threats arise from data collection choices and uneven public reporting and should be considered when generalizing the results.

D. Future Directions

Future work should address these gaps through concrete steps that are prioritized and testable:

- **Dataset standardization:** publish a versioned schema with identifiers, temporal metadata, provenance, and reproducibility indicators to enable reruns of longitudinal analyses.
- **Broaden platform coverage:** include mobile, IoT, and cloud-resident campaigns to reduce platform bias and surface cross-platform tactic reuse.
- **Countermeasure effectiveness:** design longitudinal, quasi-experimental studies of takedown and advisory events with pre-post controls to quantify real-world impact.
- **Predictive modeling:** apply temporal clustering and survival analysis to forecast tactic adoption and decay and to support anticipatory defense.
- **Collaboration:** establish data-sharing and red-teaming agreements that improve sample diversity and validation opportunities.

Addressing these directions will improve reproducibility, expand coverage, and enable earlier, more durable interventions against evolving threats while acknowledging the persistence of legacy evasion strategies.

References

[1] C. Talos, “Threat roundup 07/13 - 07/20 (emotet, trickbot, and ryuk activity),” Online, 2018, blog Post; Cisco Talos Intelligence Group. Available at: https://blog.talosintelligence.com/2018/07/threat-roundup-0713-0720.html.

[2] ——, “Gandcrab ransomware is spreading fast,” Online, 2018, blog Post; Cisco Talos Intelligence Group. Available at: https://blog.talosintelligence.com/2018/01/gandcrab.html.

[3] SentinelOne, “The maze ransomware extortion toolkit: An in-depth analysis,” Online, 2019, analysis Report; SentinelOne. Available at: https://www.sentinelone.com/anthology/maze/.

[4] L. . T. R. Labs, “Level 3 threat research reports on the mirai botnet,” Online, 2016, analysis Report; Level 3. Available at: https://www.level3.com/en/blog/level-3-threat-research-reports-on-the-mirai-botnet/.

[5] N. C. A. (NCA), “Operation cronos: International investigation disrupts lockbit ransomware group,” Online, 2024, press Release; National Crime Agency. Available at: https://www.nationalcrimeagency.gov.uk/news/operation-cronos-international-investigation-disrupts-lockbit-ransomware-group.

[6] Mandiant, “Lockbit 3.0 ransomware affiliates continue to operate,” Online, 2022, blog Post; Mandiant. Available at: https://www.mandiant.com/resources/blog/lockbit-3-0-ransomware-affiliates.

[7] CISA and FBI, “Stop ransomware: Black basta ransomware group,” Online Advisory, 2022, cybersecurity Advisory AA22-250A; CISA. Available at: https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-250a.

[8] ——, “Stop ransomware: Rhysida ransomware,” Online Advisory, 2023, cybersecurity Advisory AA23-319A; CISA. Available at: https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-319a.

[9] P. A. N. U. 42, “Ransomhub ransomware: Emerging threat in the raas landscape,” Online, 2024, threat Report; Palo Alto Networks. Available at: https://unit42.paloaltonetworks.com/ransomhub-ransomware/.

[10] A. Freund, “Backdoor in xz/liblzma leading to sshd compromise,” Online, 2024, mailing List Post; oss-security. Available at: https://www.openwall.com/lists/oss-security/2024/03/29/4.

[11] Wikipedia contributors, “2016 dyn cyberattack,” Online, 2016, wikipedia. Available at: https://en.wikipedia.org/wiki/2016_Dyn_cyberattack.

[12] SentinelOne, “Conti ransomware continues the rebrand and retool to avoid detection,” Online, 2020, blog Post; SentinelOne. Available at: https://www.sentinelone.com/blog/conti-ransomware-continues-the-rebrand-and-retool-to-avoid-detection/.

[13] Mandiant, “Darkside ransomware: A year in review,” Online, 2021, blog Post; Mandiant. Available at: https://www.mandiant.com/resources/blog/darkside-ransomware-a-year-in-review.

[14] CISA, FBI, and HHS, “Takedown of hive ransomware,” Online Advisory, 2023, cybersecurity Advisory AA22-344A; CISA. Available at: https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-344a.