SYSTEMATIC LITERATURE REVIEW ON MALICIOUS COMMAND AND CONTROL: TYPES, TECHNIQUES, TOOLS, CHALLENGES AND RESEARCH DIRECTIONS

JOSEPH A. OJENIYI, ANYIGOR CHIGBO CEPHAS, S.O.SUBAIRU, NOEL M. DOGONYARO, SULEIMAN AHMAD & ANDREWS UDUIMOH
Department of Cyber Security Science, Federal University of Technology, Minna, Nigeria
E-mail: Ojeniyija@futminna.edu.ng **Phone No:** +2347041639013

Abstract
Malicious Command and Control (C2) traffic is a critical enabler of modern cyberattacks, allowing remote management of compromised systems. This paper presents a systematic literature review (SLR) of 14 primary studies published between 2014 and 2025, identified via a structured PRISMA process across six databases (IEEE, ACM, EEE Xplore, etc.). The review synthesizes six dominant C2 models—centralized, P2P, DGA, fast-flux, cloud abuse, and encrypted traffic—and evaluates detection methods including DNS entropy and machine learning. Results indicate that while detection accuracy for P2P and DGA has improved, encrypted traffic and cloud-based C2 remain significant blind spots. We identify a critical need for explainable AI (XAI) and metadata-based analysis. This review provides a roadmap for researchers and practitioners to develop more resilient, automated threat detection frameworks.

Keywords: command-and-control traffic, C2 detection, botnet communication, and malicious traffic.

Introduction

Malicious Command and Control (C2) infrastructures are central to modern cyberattacks, enabling adversaries to remotely manage infected devices, issue malicious commands, and coordinate campaigns such as botnets and advanced persistent threats (Joseph *et al.*, 2014; Vladimir *et al.,* 2020). These infrastructures provide attackers with persistence and scalability, making them a critical target for defenders. Without C2 channels, large-scale attacks would be significantly less effective, which is why detecting and disrupting them has become a priority for cybersecurity researchers and practitioners.

Attackers continually adapt their methods to evade detection. Techniques such as domain generation algorithms, fast-flux networks, social media-based C2, and cloud service abuse have emerged as resilient alternatives to traditional centralized servers (Turki *et al.*, 2023; Cheng *et al.,* 2024). These approaches exploit trusted platforms and encrypted communication channels, blending malicious traffic with legitimate activity and complicating defensive measures. For example, Nicolas *et al.* (2016) demonstrated through the Locked Shields cyber defense exercise that communication between compromised hosts and C2 servers remains consistent across networks, making it a reliable detection target. This highlights the importance of leveraging realistic datasets in validating detection strategies, even though such datasets are scarce.

Despite extensive research, existing studies remain fragmented. Some focus narrowly on specific malware families, while others emphasize detection algorithms without integrating broader perspectives. Few works provide a holistic synthesis of C2 infrastructures across diverse contexts, leaving a gap in understanding the evolution of malicious C2 traffic and the effectiveness of defensive strategies. Fabian *et al.* (2023) and similar studies show promising machine learning approaches, yet challenges such as encrypted traffic visibility, false positives, and dataset scarcity persist.

The objective of this paper is to systematically review malicious C2 traffic research published between 2014 and 2025. Specifically, it classifies types of infrastructures, analyzes detection techniques, identifies persistent challenges, and highlights future research directions. By consolidating findings across multiple domains, this review aims to provide a unified perspective that informs both academic research and practical defense strategies, ultimately contributing to stronger resilience against evolving cyber threats.

Related Work

Gardiner *et al.,* 2014. The paper Command and Control: Understanding, Denying and Detecting' examined the critical role of Command and Control (C2) channels in advanced persistent threats (APTs) and targeted cyber attacks, highlighting how attackers establish covert communication to control compromised systems and exfiltrate sensitive data. It reviews the evolution of C2 infrastructures from centralized IRC/HTTP models to more resilient decentralized P2P and covert channels leveraging DNS, social networks, and Tor, alongside the evasive techniques attackers use to bypass signature-based detection, dynamic analysis, and reputation systems. On the defensive side, the paper outlines detection strategies such as monitoring DNS/IP traffic, identifying anomalies like periodic beaconing or fast-flux patterns, and enforcing denial measures through network segmentation, rate-limiting, and blocking unused communication mechanisms. The results emphasize that while intrusion prevention is difficult, disrupting or detecting C2 activity significantly reduces the impact of successful compromises, making C2 detection a vital line of defense against modern cyber threats.

Plohmann *et al.,* (2016). The paper 'A Comprehensive Measurement Study of Domain Generating Malware' systematically analyzes 43 malware families using domain generation algorithms (DGAs). By reverse-engineering and reimplementing these DGAs, the authors pre-computed over 18 million possible domains and studied their registration status across 9 billion WHOIS records. Results show that pre-computation enables reliable identification of malware campaigns, profiling of botmasters' registration strategies, and recognition of pitfalls in past takedown efforts. The study provides a taxonomy of DGAs, highlights common implementation flaws, and introduces DGArchive, a web service to check domains for DGA origins, thereby offering defenders a powerful tool to anticipate and disrupt botnet infrastructures.

Ghafir *et al.*, (2017) 'Botnet Command and Control Traffic Detection Challenges: A Correlation-based Solution' He addressed the growing threat of botnets and the difficulties in detecting their command-and-control (C&C) traffic. It highlights the limitations of existing intrusion detection systems and proposes a novel correlation-based approach that combines multiple detection modules—malicious IP address detection, malicious SSL certificate detection, domain flux detection, and Tor connection detection—into a collaborative voting framework. Each module independently analyzes network traffic for specific C&C techniques, and their outputs are correlated to raise alerts, thereby reducing false positives and improving accuracy. Implemented on top of the Bro network security monitor, this flexible and extensible system is designed to support real-time detection and adapt to new attack methods, offering a promising solution to the persistent challenge of botnet C&C traffic identification.

Neetesh *et al.*, (2018). The paper 'Impact Evaluation of Malicious Control Commands in Cyber-Physical Smart Grids' investigated how cyber attackers can disrupt smart grids by injecting malicious control commands into critical components such as generators, transformers, and circuit breakers. The authors develop a cyber-physical co-simulator tool for real-time security assessment, capable of detecting threats like BlackEnergy and monitoring grid health against command injection attacks. Tested on a 42-bus system with 24 substations, the tool evaluates system susceptibility, adversary access points, and threat capability, while also measuring performance factors such as scalability, robustness, and response time. The results show that

malicious commands can cause severe instability and outages, but the proposed tool effectively detects and mitigates such attacks within seconds, providing operators with situational awareness and actionable defense strategies to safeguard smart grid infrastructure.

Đorđe *et al*., (2020). The paper 'Analysis and Characterization of IoT Malware Command and Control Communication' He investigated the dynamic behavior of IoT botnets, focusing on the Mirai and Gafgyt families, which have been responsible for large-scale DDoS attacks. Using Raspberry Pi devices infected with real malware samples, the authors analyze command-and-control (C&C) communication patterns across three phases—connection establishment, maintenance, and attack execution—while comparing them to normal traffic flows. Their findings reveal that IoT botnet traffic exhibits distinctive characteristics such as periodic communication, avoidance of DNS queries, and unique statistical footprints (e.g., higher Pearson correlation, lower throughput variance) that differentiate it from legitimate traffic. The study emphasizes that these behavioral patterns can serve as reliable indicators for detection mechanisms, offering valuable insights into designing smarter intrusion detection systems to counter evolving IoT botnet threats.

Vladimir *et al.,* (2020). The paper 'Malware Command and Control Over Social Media: Towards the Server-less Infrastructure' explores emerging trends in using social media platforms as C2 infrastructure. It identifies five key techniques: hiding commands in text posts, embedding instructions via steganography in images, leveraging public cloud services as C2 servers, using domain or user generation algorithms (DGA/UGA) to create resilient accounts, and exploiting posts of public figures to conceal commands. Results show that these methods make detection extremely difficult, as they blend malicious traffic with legitimate social media activity. The authors propose that attackers could combine these trends into a stealthy, server-less C2 model, while defenders must develop new detection strategies beyond traditional domain and traffic analysis.

CyBOK, (2021) 'Malware and Attack Technologies Knowledge Area' provided a taxonomy and analysis of malware types, behaviors, and infrastructures. It categorizes malware by dimensions such as persistence, spreading method, system stack layer, and coordination (e.g., botnets). Results highlight that modern malware increasingly uses obfuscation, dynamic updates, and coordinated infrastructures to evade detection. The paper emphasizes the role of malware in cybercrime and APTs, showing how botnets exemplify noisy, large-scale attacks while APT malware operates stealthily and persistently. It concludes that effective defense requires comprehensive malware analysis techniques—static, dynamic, fuzzing, and symbolic execution—alongside awareness of the underground ecosystem that supports malware. development and monetization.

Benard *et al.,* (2022). 'C2 – Command and Control: A System of Systems to Control Complexity' he explored the theoretical foundations of Command and Control (C2) within the broader C4ISR framework, emphasizing its role in managing complexity in military and civil operations. It situates C2 as both a cybernetic and systemic process, integrating command (authority and decision-making) with control (regulation and feedback) to achieve strategic objectives in dynamic environments. The authors highlight how modern C2 has evolved from traditional battlefield practices into a technologically driven system of systems, incorporating communication, intelligence, surveillance, and decision-support tools. The results of their analysis show that C2 is not only central to military strategy but also applicable to crisis management, industrial projects, and complex organizational systems, where it reduces uncertainty, enhances situational awareness, and enables faster, more effective decision-making. Ultimately, the paper concludes that while C2 is well formalized in practice, there is

still no unified scientific theory, and its complexity requires ongoing research to integrate cybernetics, information theory, and social sciences for more robust models of control.

Almuthanna *et al.*, (2022). The paper 'EARLYCROW: Detecting APT Malware Command and Control over HTTP(S) Using Contextual Summaries' introduced EARLYCROW, a novel detection system designed to identify advanced persistent threat (APT) malware communications that mimic legitimate HTTP(S) traffic. Building on a threat model informed by tactics, techniques, and procedures (TTPs) observed in real APT campaigns, the authors propose PAIRFLOW, a multipurpose network flow format that captures contextual information such as packet behavior, DNS usage, URL structures, and time-based traffic patterns. EARLYCROW leverages these contextual summaries with random forest classifiers to distinguish malicious flows from legitimate or botnet traffic. The results show that EARLYCROW achieves strong performance, with a macro average F1-score of 93.02% and a low false positive rate of 0.74%, demonstrating its effectiveness in detecting evasive APT communications at an early stage and outperforming traditional intrusion detection approaches.

Ramos and Wang, (2023) conducted a study on "Detecting Stealthy Cobalt Strike C&C Activities via Multi-Flow based Machine Learning". This paper presents a novel machine learning–based approach for detecting stealthy Cobalt Strike command-and-control (C&C) activities hidden within encrypted HTTPS traffic, a challenge given the tool's ability to mimic legitimate traffic and evade traditional intrusion detection systems. The authors analyze real-world Beacon traffic patterns and propose multi-flow features such as session duration, periodicity, client/server data ratios, and initiation behaviors that capture inherent differences between malicious and normal HTTPS sessions. They construct group-based features across multiple TLS flows and evaluate several algorithms, finding that neural networks achieve the strongest performance with a 90.9% true positive rate and only 0.4% false positives, outperforming random forest, SVM, Naïve Bayes, and k-means. By leveraging diverse datasets including CICIDS17, CTU-Normal-20, lab-generated Beacon traffic, and real-world attack traces, the study demonstrates practical effectiveness and robustness against varied malleable C&C profiles. Overall, the paper makes a significant contribution to intrusion detection research by showing that multi-flow analysis combined with machine learning can reliably uncover stealthy Cobalt Strike activities that evade conventional detection.

Turki *et al.,* (2023) published this book 'Systematic Review Abuse of Cloud-Based and Public Legitimate Services as Command-and-Control (C&C) Infrastructure', presented a comprehensive literature review of how cyber attackers exploit trusted cloud and public services—like Dropbox, Slack, Twitter, and Gmail—as covert C&C channels to control botnets and exfiltrate data. The identifying nine distinct attack techniques including steganography, process injection, and COM hijacking. It highlights a critical gap in cybersecurity: while abuse of these platforms is rising, only a handful of studies propose effective detection methods. The authors introduce new taxonomies for attack strategies and C&C architectures, offering valuable insights into the evolving threat landscape and underscoring the urgent need for advanced detection and defense mechanisms against CPLS-based botnet operations.

Connor *et al.,* (2024). The thesis 'Command and Control Mechanisms for Post-Exploitation' investigated how attackers leverage command-and-control (C2) frameworks to maintain persistence, evade detection, and conduct malicious activities after initial exploitation of client-server systems. It analyzes C2 architectures, implant implementation methods, and evasion techniques such as obfuscation, domain generation algorithms, and endpoint detection and response (EDR) bypassing. Using case studies like the SolarWinds (SUNBURST, TEARDROP, RAINDROP) supply-chain attack and LummaC2 malware-as-a-service, the work demonstrates how modern C2 frameworks mimic legitimate traffic, employ redirectors, and utilize beaconing

to avoid detection. The results show that C2 frameworks are increasingly prevalent malicious server detections rose over 30% in 2022 and 109% in 2023—driven by open-source development and advanced evasion strategies. The thesis concludes that understanding these mechanisms is critical for designing effective defensive measures, as C2 activity remains the linchpin of persistence and data exfiltration in post-exploitation attacks.

Wang *et al.,* (2024). The paper 'Discovering Command and Control (C2) Channels on Tor and Public Networks Using Reinforcement Learning' presented a reinforcement learning (RL) approach to automatically identify resilient C2 attack pathways across both public internet and Tor-based communication channels. By modeling the attack lifecycle—infection, connection, and exfiltration the RL agent learns to bypass firewalls and optimize stealthy data transfers. Experiments using a typical enterprise network topology showed that the RL agent successfully executed attacks about 60% of the time, with 69 out of 100 simulated attack paths completing full payload exfiltration undetected. The results highlight that the agent strategically exploited intermediate nodes, used both Tor and public channels for communication, and adapted upload behaviors to avoid triggering alerts. Overall, the study demonstrates that RL can effectively automate the discovery of C2 channels, revealing both the vulnerabilities of current defenses and the potential for RL-based tools to strengthen detection and mitigation strategies.

Parssegny *et al*., (2025). The paper 'Striking Back at Cobalt: Using Network Traffic Metadata to Detect Cobalt Strike Masquerading Command and Control Channels' proposed a machine learning-based method to detect Cobalt Strike C2 traffic that mimics benign services and uses encryption. By focusing solely on network metadata (such as packet size, direction, and timing) rather than deep packet inspection, the method adapts to different malleable profiles and protocols (HTTP, HTTPS, DNS). Results show mean F1 scores between 0.78 and 1 across real-world configurations, equaling or surpassing prior approaches while remaining explainable and practical for production environments. The study demonstrates that metadata-based detection can overcome Cobalt Strike's masquerading and encryption, offering defenders a scalable and effective countermeasure.

Methodology

Search Strategy

A systematic search was conducted across six electronic databases IEEE, ScienceDirect, IEEE Xplore, ACM Digital Library, SpringerLink, and Web of Science using keywords such as command-and-control traffic, C2 detection, botnet communication, and malicious traffic. This process yielded 64 initial records.

Table 1: **Search Strategy and keywords; summarizes databases and number of articles retrieved**

S/N	Databases	No. of Articles
1	IEEE	7
2	Science Direct	15
3	EEE Xplore	12
4	ACM.Digital Library	11
5	SpringerLink	10
6	Web of Science	9
	Total	**64**

Inclusion and Exclusion Criteria

Studies published between 2014 and 2025 were included if they focused on malicious C2 traffic, were written in English, and were available in full text. Review papers, incomplete works, or studies unrelated to malicious traffic were excluded. The aim was to capture original research that proposed methodologies, detection techniques, or analysis of malicious C2 infrastructures.

Table 2: **Inclusion and Exclusion Criteria; shows rationale clearly**

S/N	Criteria	Rationale
1	Original paper; not review/survey	It should focus on Malicious Command and control traffic
2	The intended result is proposed on methodologies and techniques	The aim is to provide a comprehensive review of Malicious Command and control traffic techniques, identify common challenges faced by researchers and practitioners.
3	The paper should be a complete one (full-length)	Non-full-length paper can contain the key points but not sufficient with the required information.
4	The Research publication should be in English	The paper should be in English
5	All the publication should be from 2015-2025	This systematic literature review is in a period of 11 years, from 2014-2025

Screening/Quality Assessment

Following PRISMA guidelines, duplicates and irrelevant records were removed, resulting in 14 studies included for full analysis. These studies represent diverse approaches to detecting and mitigating malicious C2 traffic.

Following PRISMA guidelines, duplicates and irrelevant records were removed, resulting in 14 studies included for full analysis. These studies represent diverse approaches to detecting and mitigating malicious C2 traffic.

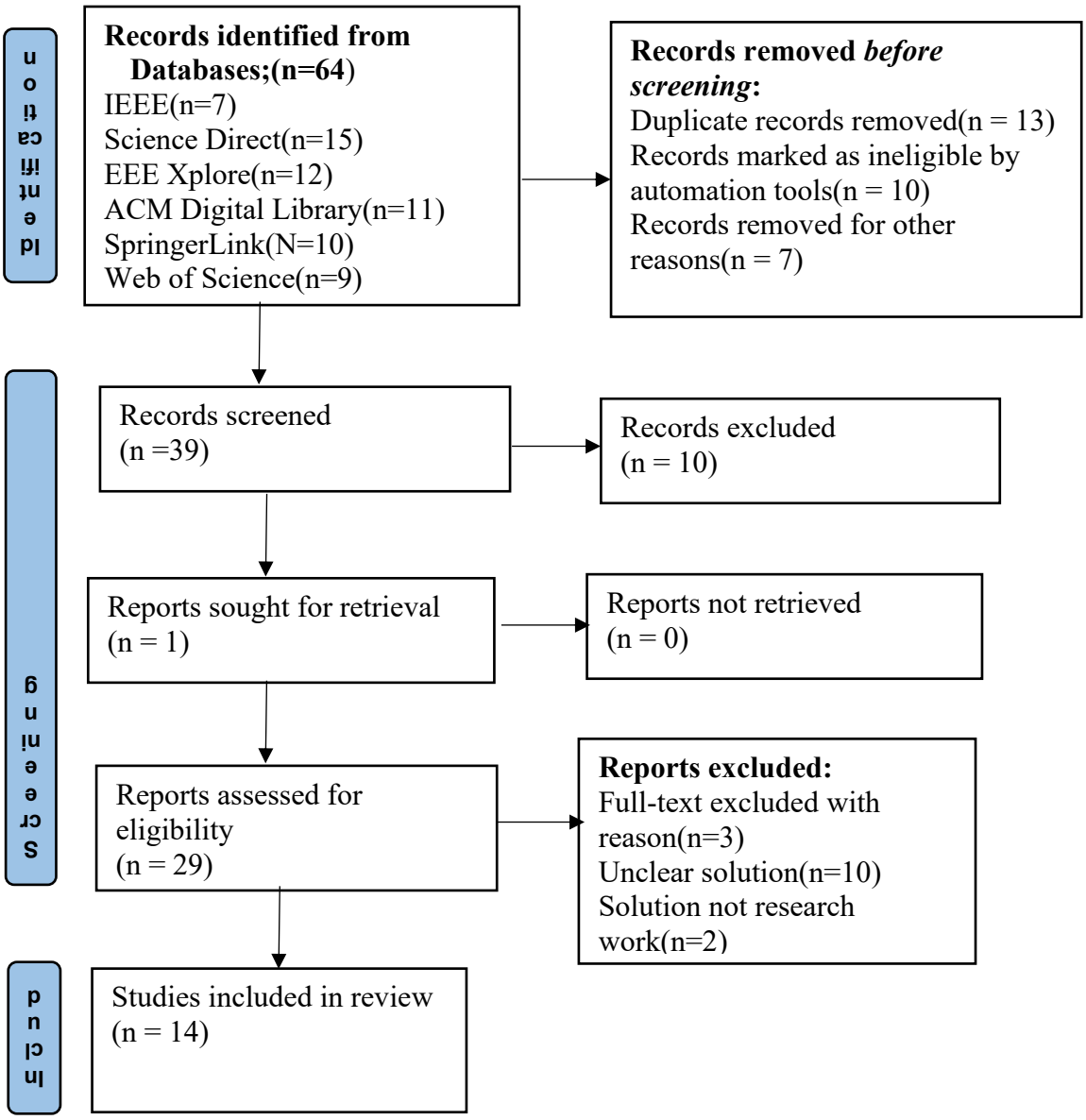

Figure 1: PRISMA DIAGRAM; this keeps the flowchart showing identification, screening, eligibility, and inclusion. This is standard in systematic reviews.

Figure 1: PRISMA DIAGRAM; this keeps the flowchart showing identification, screening, eligibility, and inclusion. This is standard in systematic reviews.

Data Extraction

For each study, information was extracted on research methodology, datasets used, detection techniques, results, and conclusions. Notably, some studies employed specialized datasets such as the Locked Shields cyber defense exercise logs (Nicolas, *et al.,* 2016), which provided rare ground-truth attacker activity. While this review did not directly analyze such datasets, their use in prior studies was considered in evaluating detection reliability.

Results

The review identified six major types of malicious C2 traffic: centralized, peer-to-peer, domain generation algorithms, fast-flux, cloud/legitimate service abuse, and encrypted traffic. Each type was associated with specific detection challenges and tools.

Locked Shields Dataset: Nicolas *et al.* (2016) demonstrated that communication between compromised hosts and C2 servers is consistent across networks, making it a reliable detection target. This dataset provided rare ground-truth attacker logs, which were used to validate detection methods.

"The study uses data from Locked Shields, the world's largest cyber defense exercise, which simulates real-world conditions by testing defenders against severe attacks while providing rare ground-truth logs of attacker activity" (Nicolas *et al*., 2016).

Machine Learning Approaches: Ramos *et al.* (2023) showed that multi-flow neural networks achieved a 90.9% true positive rate in detecting stealthy Cobalt Strike C2 traffic.

"Neural networks achieve the strongest performance with a 90.9% true positive rate and only 0.4% false positives" Ramos *et al.* (2023)

Cloud Service Abuse: Turki *et al.* (2023) highlighted how attackers exploit Dropbox, Slack, and Twitter as covert C2 channels, noting limited detection strategies.

"Cyber attackers exploit trusted cloud and public services—like Dropbox, Slack, Twitter, and Gmail as covert C&C channels" (Turki *et al*., 2023).

Reinforcement Learning: Wang *et al*. (2024) introduced RL agents capable of discovering resilient C2 channels across Tor and public networks, with 60% success in undetected payload exfiltration.

"Experiments showed the RL agent successfully executed attacks about 60% of the time, with 69 out of 100 simulated attack paths completing full payload exfiltration undetected" Wang *et al*. (2024)

The distribution of these publications by year is as follows:

2014: 1 paper	2015: nil	2016: 1 paper	2017: 1 paper
2018: 1 paper	2019: nil	2020: 2 papers	2021: 1 paper
2022: 2 papers	2023: 2 papers	2024: 1 paper	2025: 1 papers

(as shown in Figure 2)

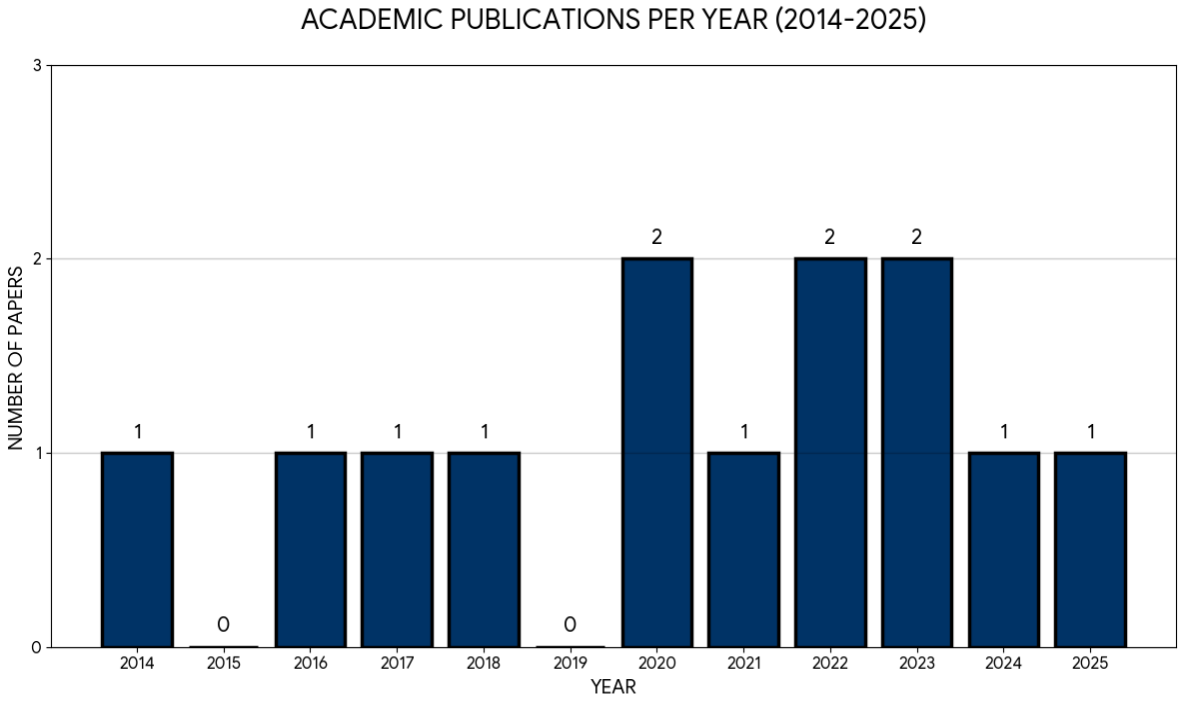

Figure 2: BAR CHART; visually shows research trends over time.

Types, Techniques, Tools, Challenges, and Research Directions

Types of Malicious C2 Traffic

Centralized C2: Single server architecture; easily disrupted but still used in simple malware.
Peer-to-Peer (P2P) C2: Decentralized communication among bots; resilient against takedowns.
Domain Generation Algorithms (DGAs): Automated domain creation complicates blacklisting.
Fast-Flux Networks: Rapid IP address rotation hides malicious servers.
Cloud/Legitimate Service Abuse: Exploiting cloud platforms, APIs, and messaging services.
Encrypted C2 Traffic: TLS/HTTPS adoption conceals payloads.

Table 3: Comparative Table of C2 Models, Techniques & Tools

S/N	Technique /Model	Author(s)/ Year	Description	Key Insights/Risks	Practical Tools
1	Centralized C2	Bernard Claverie & Gilles Desclaux (2016, 2022)	Traditional hierarchical model where authority and decision-making are concentrated at the top.	Works in stable environments but becomes a bottleneck in complex, dynamic operations. Slows decision-making, risks overload at the center, and lacks adaptability.	SIEM platforms (Splunk,IBM QRadar),Decision-support dashboards, Command simulation tools.
2	Peer-to-Peer (P2P) C2	SpecterOps (2019); GitHub frameworks(2020s)	Distributed network where compromised hosts communicate directly with each other.	Resilient, no single point of failure. Hard to detect since traffic resembles normal peer connections. Used in malware like Storm and GameOver Zeus.	IDS/IPS (Snort, Suricata), NetFlow analyzers, Zeek (Bro) for traffic inspection.

3	Domain Generation Algorithms(DGAs)	Conficker (2008); Zscaler (2024)	Malware generates thousands of pseudo-random domains daily to locate active C2servers.	Highly evasive: defenders cannot block all domains. Example: Conficker generated 50,000/day. Requires DNS anomaly detection or ML-based defenses.	Passive DNS monitoring, Machine learning classifiers (TensorFlow, Scikit-learn), OpenDNS Umbrella.
4	Fast-Flux Networks	CISA(2025); FBI advisories	Rapidly changing DNS records and botnets act as reverse proxies to hide trueC2 servers.	Dynamic obfuscation: IP addresses change constantly. Common in phishing and malware distribution. Difficult to dismantle.	DNS anomaly detectors, Threat intelligence feeds, Wireshark for packet analysis.
5	Cloud/Legitimate Service Abuse	MDPI (2023); Unit 42(2025)	Attackers exploit trusted cloud services (Dropbox, Google Drive, AWS) as covert C2channels.	Blends with legitimate traffic. Hard to block without disrupting business. Example: AWS X-Ray repurposed for hidden C2.	CASB (Cloud Access Security Brokers),Cloud-native monitoring (AWS GuardDuty, Azure Sentinel),API activity logging.
6	Encrypted C2 Traffic	Palo Alto Networks (2025); Zscaler (2024)	Attackers encrypt C2 communications (HTTPS/TLS) to blend with normal traffic.	Hard to detect since security tools see only encrypted streams. Increasingly common with frameworks like Sliver. Requires behavioral analysis or deep packet inspection.	SSL/TLS inspection tools, Deep Packet Inspection (DPI), Behavioral analytics (Darktrace, Vectra AI).

Challenges

Encrypted Traffic: TLS hampers payload inspection because defenders cannot see inside encrypted streams. As Claverie, *et al.*, 2022 note, control depends on 'feedback and regulation', and encryption reduces visibility, weakening situational awareness.

False Positives: Legitimate cloud traffic often resembles malicious C2. MDPI (2023) emphasize that cloud-native traffic patterns blur the line between benign and malicious, complicating detection.

Adaptive Adversaries: Attackers evolve evasion techniques rapidly. Zscaler, 2024 highlight that DGAs and adaptive malware families constantly change signatures, forcing defenders to recalibrate.

Dataset Scarcity: Limited labeled datasets hinder ML training. Claverie, *et al.*, 2016 argue that the absence of a unified scientific theory of C2 mirrors the lack of robust datasets for training detection models.

Scalability: Large networks impose high computational costs. CISA (2025) warn that fast-flux detection at scale requires significant computational resources, especially in enterprise networks.

Evasion Techniques: Steganography, covert channels, protocol tunneling hide malicious C2. Palo Alto Networks (2025) document how modern adversaries embed C2 in covert channels, making detection extremely difficult.

Research Directions

Explainable AI (XAI): Improve interpretability of ML models. Claverie & Desclaux (2022) call for integrating 'cybernetics and social sciences', which aligns with XAI's goal of transparent decision-making.

Cloud-Native Defenses: Tailored detection for SaaS and cloud platforms. Unit 42 (2025) show how attackers abuse AWS services, underscoring the need for cloud-native defenses.

Encrypted Traffic Analysis: Use metadata (timing, flow statistics) instead of payload inspection. Zscaler (2024) recommend metadata-based detection for TLS traffic, echoing Claverie & Desclaux's emphasis on feedback loops.

Federated Learning: Collaborative ML approaches preserving privacy. MDPI (2023) propose federated learning as a way to share intelligence across organizations without exposing raw data. Cross-Layer Detection: Integrating host-based and network-based signals. CISA (2025) advise combining host telemetry with network monitoring to detect fast-flux and P2P C2.

Automated Threat Intelligence Integration: Real-time IoC sharing. Palo Alto Networks (2025) emphasize automated IoC feeds as critical for reducing uncertainty and enhancing situational awareness.

Table 4: Comparative Table of C2 Models, Techniques, Challenges & Directions

S/N	Technique/ Model	Detection Techniques (Claverie & Desclaux,2016;2022)	Challenges	Research Directions
1	Centralized C2	Systemic monitoring &feedback loops: detect overload at the center by tracking decision latency and situational awareness gaps.	Scalability (CISA,2025); Dataset scarcity (Claverie & Desclaux,2016).	Explainable AI (Claverie & Desclaux,2022); Cross-layer detection (CISA,2025).
2	Peer-to-Peer (P2P)C2	Network behavior analysis: detect unusual peer-to-peer traffic patterns.	Adaptive adversaries (Zscaler,2024).	Federated learning (MDPI,2023); Automated threat intelligence (Palo Alto,2025).
3	Domain Generation Algorithms (DGAs)	Information-theoretic anomaly detection: monitor DNS queries for entropy and randomness	Dataset scarcity (Claverie & Desclaux,2016).	Explainable AI (Claverie & Desclaux, 2022);Metadata-based encrypted traffic analysis (Zscaler,2024).
4	Fast-Flux Networks	Feedback-driven monitoring: track rapid IP changes and short TTL values.	Scalability (CISA,2025).	Cross-layer detection (CISA,2025).
5	Cloud/Legitimate Service Abuse	Contextual situational awareness: monitor abnormal usage of cloud APIs.	False positives (MDPI, 2023).	Cloud-native defenses (Unit 42,2025); Automated IoC sharing (Palo Alto,2025).

6	Encrypted C2 Traffic	Behavioral analysis & systemic feedback: detect anomalies in encrypted traffic volume, timing, and endpoint behavior.	Encrypted traffic hampers payload inspection (Claverie & Desclaux,2022).	Encrypted traffic analysis via metadata (Zscaler,2024); Federated learning (MDPI,2023).

Conclusion

This systematic literature review successfully achieved its objective of classifying malicious Command and Control (C2) infrastructures and evaluating the efficacy of current detection methodologies published between 2014 and 2025. By synthesizing 14 primary studies, the research identified six dominant C2 models: centralized, P2P, DGA, fast-flux, cloud service abuse, and encrypted traffic. The findings demonstrate a critical shift in the threat landscape; while traditional DNS-based detection remains effective for older models like DGAs, multi-flow neural networks and metadata-based analysis are now essential for uncovering stealthy, encrypted channels such as Cobalt Strike, with some models achieving a 90.9% true positive rate. This review's unique contribution lies in providing a performance-based comparison of these AI-driven detection tools against specific C2 evasion techniques, offering a clearer roadmap than existing fragmented surveys.

The scope of this review is limited by its reliance on a small sample of 14 high-quality papers that met the strict inclusion criteria. Additionally, the study is restricted to English-language databases, which may omit relevant regional research. Furthermore, because the C2 landscape evolves rapidly, the findings primarily reflect established trends up to early 2025 and may not fully account for emerging "zero-day" infrastructures.

For cybersecurity practitioners, these results underscore the need to move beyond payload inspection rendered increasingly obsolete by encryption and adopt metadata-based flow analysis and cloud-native monitoring to distinguish malicious traffic from legitimate business activity. Future research should prioritize Explainable AI (XAI) to improve the transparency of automated detection and Federated Learning to facilitate cross-organizational intelligence sharing without compromising privacy. Finally, there is an urgent need to contextualize these global trends within the Nigerian digital landscape, specifically focusing on localized threats to fintech and telecommunications infrastructure.

References

Joseph, G., Marco, C., & Shishir, N. (2014) Command and control; Understanding, denying and detecting. https://arxiv.org/abs/1408.1136.

Daniel, P., Fraunhofer, F., Khaled, Y., Michael, K., Johannes, B., Elmar, G., & Fraunhofer K. (2016) A comprehensive measurement study of domain generating malware. https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_plohmann.pdf

Ghafir, I., & Prenosil, V., & Hammoudeh, M. (2017). Botnet command and control traffic detection challenges: A correlation-based solution. *International Journal of Advances in Computer Networks and Its Security– IJCNS,* 7, 2250-3757. https://www.researchgate.net/publication/315843154.

Saxena, N., Xiong, L., Chukwuka, V. & Grijalva, & (2021). Impact Evaluation of Malicious Control Commands in Cyber-Physical Smart Grids, in IEEE Transactions on Sustainable Computing, 6,(2), 208-220, doi: 10.1109/TSUSC.2018.2879670. https://ieeexplore.ieee.org/document/8523803.

Đorđe, D. J., & Pavle, V. V. (2020) Analysis and characterization of IoT malware command and control communication. https://journal.telfor.rs/Published/Vol12No2/Vol12No2_A2.pdf.

Radunović V., & Veinović M. (2020). Malware command and control over social media: Towards the server-less infrastructure. https://doi.org/10.2298/sjee2003357.

CyBOK (2021). Malware and attack technologies knowledge area https://cybok.org/media/downloads/Malware_Attack_Technologies_v1.0.1.pdf?.

Claverie, B., & Desclaux, G. (2022). C2 - Command and control: A system of systems to control complexity. https://www.researchgate.net/publication/363583535

Fabian, M. R., & Xinyuan, W. (2023). Detecting stealthy cobalt strike C&C activities via multi-flow based machine learning. https://people.cs.gmu.edu/~xwangc/Publications/ICMLA2023-CobaltStrikeMFMLDetect.pdf.

Hummel, C. (2024). Command and control mechanisms for post-exploitation. LSU Master's Theses. 6043. https://repository.lsu.edu/gradschool_theses/6043.

Cheng, W., Christopher, R., Abdul Rahman, R. C., Daniel, R., Tyler, C., Dhruv, N., & Edward, B. (2024). Discovering command and control (C2) channels on tor and public networks using reinforcement learning. https://arxiv.org/abs/2402.09200.

Clément, P., Johan, M., Olivier, L., & Pierre, C. (2025). Striking back at cobalt: Using network traffic metadata to detect cobalt strike masquerading command and control channels". https://arxiv.org/abs/2506.08922.