Literature Review on C2 Traffic Detection Using Machine Learning and Network Metadata

A Comprehensive Summary of Recent Advances in Encrypted C2 Detection

Prepared for Research Purposes

July 12, 2025

This document summarizes key studies on detecting Command and Control (C2) traffic using machine learning and network metadata, with a focus on encrypted traffic and evasion scenarios.

Contents

1 Literature Review

This section reviews pivotal studies on detecting Command and Control (C2) traffic using network metadata and machine learning, emphasizing encrypted traffic and evasion scenarios. The works inform a methodology for detecting C2 traffic with lightweight models trained on synthetic data.

1.1 Parssegny et al. (2025)

Parssegny et al. proposed an ML-based method for detecting Cobalt Strike C2 traffic using packet sizes, timing, and flow direction [**?**]. Their dynamic model adapts to traffic patterns, offering explainability and real-world deployability. Limited to Cobalt Strike, it inspired broader C2 framework evaluation and evasion testing in synthetic traffic.

1.2 Detection of Malicious DNS-over-HTTPS Traffic (2023)

An unnamed 2023 study used an autoencoder for anomaly detection in DNS-over-HTTPS (DoH) traffic, achieving a 99% F1-score for zero-day threats [**?**]. Focused on DoH, it aligns with applying lightweight ML to HTTP and DNS tunneling scenarios for C2 detection.

1.3 Barradas et al. (2024)

Barradas et al. explored C2 detection in TLS 1.3 using certificate size and protocol behavior, achieving over 93% detection rates [**?**]. Lacking full flow metadata and evasion testing, their work supports incorporating richer metadata and simulating evasion tactics.

1.4 Qing et al. (2023)

Qing et al.'s RAPIER framework detects encrypted malicious traffic despite label noise, using GANs and autoencoders [**?**]. Applied to general malware, it informs C2-specific detection with synthetic datasets and evasion strategies like jittered timing.

1.5 Wang & Thing (2023)

Wang & Thing introduced "Enc Feature" for encrypted traffic, achieving a 99.72% F1-score with deep learning [**?**]. Their feature engineering supports a metadata-driven approach, complemented by simpler, interpretable models like Decision Trees.

1.6 Ramos & Wang (2022)

Ramos & Wang detected stealthy Cobalt Strike traffic using timing features, with Random Forest achieving a 50% true positive rate [**?**]. Limited profiles prompted improved accuracy and diverse evasion simulations.

1.7 Patel et al. (2023)

Patel et al. used ML for malware detection in TLS traffic, leveraging handshake details and packet sizes [**?**]. Focused on general malware, it supports narrowing to C2 callback detection under evasion conditions.

1.8 Waqas (2024)

Waqas reviewed DoH misuse for C2, noting Random Forest's 99.89% accuracy [**?**]. As a survey, it lacks implementation, but it validates practical C2 detection over DoH.

1.9 Sebakara & Jonathan (2025)

Sebakara & Jonathan's privacy-preserving ML detected RATs with 75% accuracy using metadata [**?**]. Focused on RATs, it informs synthetic C2 traffic generation for improved accuracy and evasion resistance.

1.10 Yang et al. (2024)

Yang et al.'s TrafCL used contrastive learning for robust detection [**?**]. Its deep learning reliance prompts simpler supervised models for C2-specific detection under evasion.

1.11 Xavier et al. (2023)

Xavier et al. showed Metasploit C2 detection (99% accuracy) but vulnerability to minor modifications [**?**]. This inspired evasion simulations to test lightweight ML model resilience.

1.12 Alwhbi et al. (2024)

Alwhbi et al.'s survey validated ML techniques for encrypted traffic analysis [**?**]. It supports applying transfer learning and lightweight models to C2 detection with evasion scenarios.

1.13 AlShaikh et al. (2023)

AlShaikh et al. achieved 99.98% accuracy for IoT C2 detection [**?**]. Limited to known botnets, it informs simulating diverse C2 behaviors under adversarial conditions.

1.14 Alageel & Maffeis (2025)

Alageel & Maffeis's EARLYCROW achieved a 93% F1-score for APT C2 over HTTP(S) [**?**]. Its feature selection supports metadata-driven detection with simpler models and DNS tunneling.

1.15 Jeong et al. (2024)

Jeong et al.'s CYE framework simulates C2 scenarios with reinforcement learning [**?**]. Used to generate labeled traffic, it supports lightweight ML classification and evasion testing.

1.16 Hulsebosch (2024)

Hulsebosch explored HTTP probing for C2 server fingerprinting [**?**]. Passive metadata extraction aligns with HTTP and DNS-based C2 detection.

1.17 Jain et al. (2023)

Jain et al.'s C2Store provides threat intelligence but lacks ML detection [**?**]. It informs real-time C2 detection with evasion evaluation.

1.18 Mali et al. (2025)

Mali et al. compared ML models for encrypted traffic, with Random Forest showing strong results [**?**]. This supports C2-focused detection with evasion scenarios.

1.19 Elmaghraby et al. (2024)

Elmaghraby et al. combined neural networks and ML for 96.8% accuracy in encrypted traffic classification [**?**]. This validates hybrid ML for C2 detection.

1.20 Hariharan (2025)

Hariharan's real-time fingerprinting used lightweight ML for encrypted threats [**?**]. This aligns with optimizing models for C2 protocols in resource-constrained environments.