Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140
Volume 14 Issue 02, 2025

Agent-vs-Agent Cyber Warfare: Autonomous AI Systems Defending Against AI-Enabled APTs

Dr Salman Arafath Mohammed
Assistant Professor, College Of Engineering, King Khalid University, Aseer Region, ABHA, KSA

Peer Review Information

Submission: 05 Oct 2025

Revision: 20 Oct 2025

Acceptance: 10 Nov 2025

Keywords

Autonomous Agents, Cyber Warfare, AI-Enabled APTs, Reinforcement Learning, Multi-Agent Systems, Cybersecurity Automation, LLM Security Defense, Threat Intelligence.

Abstract

The cyber-security ecosystem is evolving very fast, with Artificial Intelligence (AI) giving rise to both highly defensive and more sophisticated forms of Advanced Persistent Threats (APTs). AI-powered APTs are a new breed of intelligent, adaptive and self-learning cyber attackers that can autonomously use vulnerabilities, evade detection and continue operating within networks. Organizations in their turn are moving towards the shift between stationary, rule-based control and fully autonomous defensive agents able to conduct continuous monitoring, predict the threat, interrupt the attack real-time, and actively respond. It is this paper that examines the new paradigm of Agent-vs-Agent Cyber Warfare, where autonomous AI defenses indirectly respond to AI-driven APTs on dynamic digital platforms. We describe the architecture of the offensive APT agents based on AI, analyze defensive multi-agent systems (MAS), and suggest a proactive cyber-battlefield model, based on reinforcement learning (RL), large language models (LLM), and self-evolving threat intelligence. Lastly, we outline constraints, ethical aspects, and the way forward with regard to obtaining digital ecosystems in an era of autonomous cyber warfare.

Introduction

Cyber warfare has changed the human-centered decision-making to high-speed autonomous machine actions controlled by artificial intelligence. Conventional cyber defense has been based heavily on the human intelligence and signature based systems that are unable to cope with emerging threats. In meantime, state-paid and financially driven opponents are progressively turning to automation, machine learning, and generative AI to initiate multi-tier, sustained assaults that grow beyond their capacity to control.

The development of APTs facilitated by AI is a turning point in the threat landscape. These attackers use reinforcement learning, deep neural networks, autonomous reconnaissance, and natural language-enabled social engineering to penetrate and have footholds in networks. In contrast to classical malware, APTs powered by AI adapt dynamically on each step of kill chain scanning, exploitation, subsequent lateral movement, privilege escalation and exfiltration.

The response of cyber defense strategies to these adaptive threats should be within an autonomous, intelligent agent-oriented system that is able to make decisions quicker than the attackers themselves, learn the previous attacks and forecast the future adversarial actions. The article is about an Agent-vs-Agent Cyber Warfare paradigm, in which both the attack and the defense are directed by autonomous AI agents who are in a steady strategic engagement.

© 2025 The Authors. Published by MRI INDIA.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

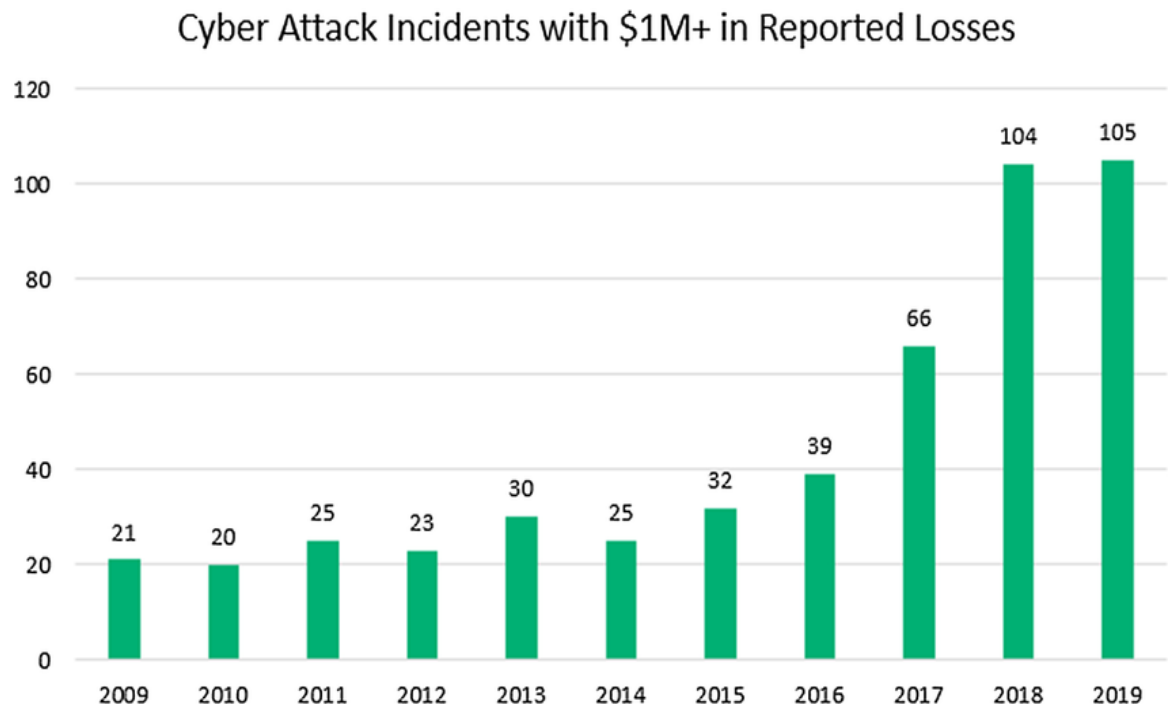

Figure 1. Relative frequency and operational emphasis of different AI-enabled APT capabilities.

Background and Related Work

An overview of evolving threats on the Internet (Advanced Persistent Threats, or APTs) is provided

1. Advanced Persistent Threats (APT) Evolution.

The last decade has witnessed a great evolution of Advanced Persistent Threats (APTs). Historically, the traditional APT groups like APT28, APT29 and Lazarus used exploits that were manually crafted and human directed operations and manually reconstructed reconnaissance to penetrate target networks. These missions involved the use of trained operators and a lot of planning in order to attain persistence and data exfiltration. Heavy AI and machine learning models have also led to the introduction of automated intelligence within the frameworks of APTs. Current APTs have turned to AI to help automate the reconnaissance, create exploits, avoid anomaly detection systems and create polymorphic behaviors that continually alter malware signatures, increasing its stealth and operational efficiency. This reconfiguration of human-reliant operations in favor of AI-supportive campaigns has more radically accelerated the pace, scale and complexity of computer attacks.

2. Rise of Autonomous AI Agents

The autonomous AI agents have become a change agent in the field of cybersecurity, utilizing the perception, learning, and decision-making systems to dynamically engage with digital space. In contrast with the previous automation scripts, which were simple, predetermined procedures, contemporary agent-based models are intelligent and can identify anomalies and risks and take complex actions without any human intervention. These agents are constantly processing environmental data, changing their strategies based on the perceived threats, and optimizing their own actions to achieve their highest goals like intrusion success or reducing threats. The development of these autonomous systems is indicative of a larger curl in cybersecurity of intelligent, self-adaptive technologies that can make decisions that react to the dynamically shifting threat environments without any direct human intervention.

3. Multi-Agent Systems (MAS) and AI with regard to Offensive Cyber Capabilities.

Multi-Agent Systems (MAS) can be used to enable greater cooperation between multiple AI agents to improve the cybersecurity activities. The defensive, monitoring, and deception agents on MAS frameworks are synchronized to identify the intrusions, predict the attackers and countermeasures. Learning (Reinforcement) in such structures enables the system to improve the defenses dynamically as the attackers vary their strategies, and this is an evolving protective network that can undergo continuous learning. On the offensive front, the enemies also use AI to enhance their operational effectiveness. Machine learning techniques help in profiling target, vulnerability detection, and

social engineering campaign automation. Auto generative AI allows attackers to generate sophisticated, highly person-centric attacks in seconds, including automated phishing messages, bogus credentials, cyber scripts, and intelligent variants of malware programs. The fusion of MAS defense and AI-based automation offense is the reason to think of the escalating arms race between attackers and defenders of the cyber world and the necessity of sophisticated, intelligent, and flexible security measures.

The AI-Enabled APT Threat Landscape

AI-enabled APTs represent a next-generation category of threats with unique capabilities:

1. Observer Unmanned Aerial Vehicles.

Contemporary independent attack agents perform very advanced reconnaissance with the help of natural language processing (NLP), machine learning (ML), and massive data mining. These advanced persistent threat (APT) systems actively gather and process open-source intelligence on scale, through tremendous digital ecosystems, on social media, uncovered databases and public repositories as well as leaked credentials. These agents are used to determine the possible vulnerabilities of a target infrastructure using predictive analytics to correlate software versions, misconfigurations, and past breach history. They also rank attack surfaces by the value, accessibility and potential impact of high-value assets to allow strategically selecting the targets without human involvement. ML enables these reconnaissance systems to constantly increase the accuracy of their estimations, enhance their predictive capability, and respond to the changing target environment security position.

2. Exploit development AI-based, Evasion, and Lateral Movement.

Generative models with AI-based offensive frameworks have changed the exploit generation cycle by being able to generate, alter and obfuscate malicious code at scale. Generative AI agents may be used to create complex adventures, conceal malicious code, as well as produce polymorphic variations that decrease the chances of them being detected by signatures. Reinforcement learning (RL) is important as it allows agents to automatically test these payloads in simulated cyber ranges prior to implementation in real networks with high success rates and low operational risks.

At the same time, real-time evasion is achievable because AI-controlled malware will constantly scan the defensive mechanisms, including firewalls, intrusion detection systems, and endpoint protection tools. The malware, according to these observations, is autonomously able to modify its behavior, the protocols it uses to communicate and binary signatures such that it has highly dynamic polymorphism to be able to stay invisible even over long periods of time. The decision engines moving as a result of RL inside the network control adaptive movement along a lateral direction. These agents construct optimal sequence of privilege escalations based on probabilistic reward functions, which strive to access as much as possible and be exposed as little as possible. This contributes to making the lateral movement more tactical, covert, and evasive to conventional monitoring systems.

3. Smart Discipline and Self-Generative Renewal.

Sophisticated APT ecosystems integrate self-healing that allows them to persist in the long term in destabilized environments. Rather than depending on one backdoor, these agents build on top of the system in multiple layers with redundant distributed footholds so that the attack can persist when some of the channels are neutralized. In case command-and-control (C2) communications are disrupted, AI-driven reasoning, commonly informed by large language models (LLMs) is used to rebuild and redesign new channels of communication with alternative protocols, backdoors, encrypted channels, or switchpoints within the network. Such self-restoration is a biological resilience, and it enables the malware infrastructure to get back on its feet following defensive intervention and continue operating.

These agents act as completely autonomous cyber adversaries with the capacity to run a long-term, stealth-based campaign that defies the conventional defense strategies through the use of dynamic adaptation, persistence reinforcement, and long-term self-improvement.

Table 1. Interaction cycle and behavioral differences between offensive APT agents and defensive autonomous agents within the proposed cyber battlefield model.

Interaction Stage	Offensive Agent Behavior	Defensive Agent Behavior
Observation	Scans environment, collects OSINT, identifies misconfigurations	Monitors telemetry, logs, user behavior, network flows
Prediction	Predicts optimal attack paths via RL	Predicts likely attacker movement or escalation strategy
Action	Deploys exploits, moves laterally, evades detection	Isolates system, deploys countermeasures, deception
Reward/Penalty	Reward for persistence, stealth, data exfiltration	Reward for detection accuracy, containment speed
Adaptation	Generates new payload variants, rebuilds C2 channels	Updates models, reduces false positives, improves response logic

Defensive Autonomous AI Systems

1. Architecture of an AI Defense Agent

The structure of a contemporary AI-driven defense agent is made to look like a layered system to offer end-to-end autonomous cybersecurity. The sensor layer at the base collects the data of various sources such as network monitors, endpoint detection and response (EDR) logs, security information and event management (SIEM) events, and user behavior analytics to offer an overall picture of network traffic. The perception layer consists of machine learning models that identify anomalies and correlate the threat indicators of these inputs to allow early identification of potential threats. The cognition layer applies large language models (LLM) and reasoning by reinforcement learning to understand these threats and their severity and decide upon the best action. Action layer executes automated protection functions like isolating systems that are attacked, patching, or implementing tricks to deceive attackers. Lastly, the learning layer makes sure that new attack samples are always adding to the model, and thus, the system is constantly adjusted to new threats.

Table 2. Capability comparison between AI-enabled offensive APT agents and autonomous defensive agents.

Offensive AI-APTs	Defensive Autonomous Agents
Autonomous reconnaissance using NLP & ML	Real-time monitoring across endpoints & networks
Polymorphic exploit generation (Generative AI)	Machine-learning-driven anomaly detection
Adaptive evasion using RL & signature mutation	Adaptive response selection (RL-based)
Self-healing persistence & redundant footholds	Self-learning threat models & continuous retraining
Covert C2 channels generated dynamically	Automated deception: honeypots & honey-tokens
Predictive vulnerability targeting	Predictive risk scoring & proactive mitigation

2. Multi-Agent Defensive Framework

Multi-agent defensive framework (MAS) is the group of specific agents providing cooperation to enhance cybersecurity defense. Detection agents are charged with the role of detecting anomalies and possible security violations. Prediction agents can decide on the possible attack paths and predict adversarial actions by analyzing patterns. Response agents undertake the necessary countermeasures such as quarantining of affected systems or counteracting threats. Deception agents can be used to increase security by using honeypots and honey tokens to redirect attackers and obtain intelligence. In the meantime, the forensic agents keep expert records that record the activity of the system and prepare reports on the occurrence of the incident to review and analyze it. A combination of these agents can be used as a synchronized defense system to offer proactive, adaptive, and intelligent security to complicated network settings.

Figure 2. Effectiveness of autonomous defensive agents across core cybersecurity protection domains.

Proposed Agent-vs-Agent Cyber Battlefield Model

This model conceptualizes cyber warfare as continuous dynamic interaction between two types of agents: Offensive AI-APTs and Defensive Autonomous Agents.

1. Environment

The cyber environment includes endpoints, networks, cloud infrastructure, identity systems, and communication channels.

2. Offensive Agent Behavior

- RL-driven decision making
- Autonomous scanning and exploitation
- Polymorphic payload selection
- Covert C2 channels using LLM-generated encryption patterns
- Continuous adaptation

3. Defensive Agent Behavior

- Continuous monitoring
- Proactive threat hunting
- Adaptive response selection
- Deception, misdirection, and decoy deployment
- Predictive risk scoring

4. Agent Interaction Dynamics

Both agents learn from each other's strategies. Interaction cycles include:

- Observation
- Prediction
- Action
- Reward/Penalty

The system evolves into a self-learning cyber ecosystem where adversarial agents and defenders co-adapt.

Methodology and Framework

Conducting training in offensive and defensive agents in cybersecurity presupposes complex approaches and specialized conditions to make sure that the latter are capable of working in the real environment. Simulated cyber ranges like Cyber Battle Sim, Capture the Flag (CTF) laboratories and custom reinforcement learning environments are often used to train the offensive agents that are designed to simulate attacks and determine the vulnerabilities of the system. These simulators offer regulated yet natural environments wherein adversarial agents may educate and perfect penetration strategies, exploit development and attack methods without putting real systems at risk. Defensive agents, conversely, are made to identify, alleviate and react to cyber threats. Their training includes attack simulation that can replicate the behavior of intrusion in the real world, historical incident information that replicate the tendency of precedent intrusion, and telemetry taken on functioning networks to comprehend the conduct of the systems under attack circumstances. Also, adversarial reinforcement learning is utilized to enable

defensive agents to change within changing attack patterns and be more responsive and receptive to new threats. At this point, the effectiveness of these agents is evaluated in terms of both quantitative and qualitative measures. Their detection rate is a measure of their capability to detect threats correctly and response latency is a measure of how quickly they respond to a threat. Other important measurements are attack disruption time, the speed with which an attack may be contained, false positive rate to determine accuracy, resource overhead to determine computational efficiency and adaptability score which determines how the agent can handle unexpected or advanced attack techniques. Through these training and evaluation strategies, cybersecurity systems will be able to train robust intelligent agents that can both initiate controlled offensive attacks and protect the complex network environments very efficiently and reliably.

Table 3: Offensive vs Defensive Agent Capabilities

Offensive AI-APT Capabilities	Defensive Autonomous Agent Capabilities
Autonomous Reconnaissance	Real-Time Monitoring
Polymorphic Exploitation	Anomaly Detection
Adaptive Evasion	Adaptive Response
Self-Healing Persistence	Self-Learning Models

Case Scenarios

AI-driven scenarios of attacks describe how current cyber war is getting more complex. The operations of phishing through AI-based spear-phishing messages are used to create countermeasures with which large language models can be utilized to create highly personalized phishing messages that are able to hazard traditional filters, and the defensive agents are able to analyze language patterns, detect abnormalities, and filter malicious communications before reaching the target. Likewise, autonomous ransomware uses dynamic encryption algorithms that dynamically vary so that these algorithms cannot be detected through a static technique, but a reinforcement learning-based protective agent can efficiently isolate the infected endpoint in a matter of seconds and, thus, lateral movements and further propagation are prevented. Attackers can use offensive agent in cloud settings to exploit improperly configured IAM policies to escalate privileges and gain access to sensitive assets, and defensive autonomous agent will (proactively) estimate the viable paths of escalation and automatically repurpose risky configurations afore they are exploited. All these situations are evidence of the dynamic nature of offensive and defensive AI agents used in the real-time cyber conflict.

Challenges and Limitations

Model drift is another major problem in autonomous cyber systems since attacker and defender models naturally degenerate with time unless retrained and updated on a continuous basis. Such decay decreases the accuracy of detection, deteriorates predictive power and, eventually, permits advanced APTs powered by AI to utilize obsolete patterns. Besides this, the issue of data privacy and security arises since successful training of defensive agents needs significant amounts of telemetry and behavioral data, which could accidentally leak sensitive or controlled information unless adequately secured. The other important concern is the explainability of autonomous decisions; in most cases these systems are based on complex reinforcement learning and deep neural networks and therefore their behaviors should be understandable to comply with legal, compliance and governance standards. Finally, multiple interacting agent real-time inference requires very large computational resources, which introduce operational overhead and can simply be deployed in low-resource or low-latensity environments. All these aspects point to the practical and technical constraints that have to be dealt with so that the agent-based cyber defense systems can be reliable and safe.

Ethical and Legal Considerations

The emergence of the autonomous actors in cyber war presents a complicated range of ethical, legal, and governance issues. Among the main issues that are to be mentioned is the danger of uncontrolled escalation. Once the attackers and the defensive side use self learning agents that have the ability of making decisions on their own, misidentification or overreaction can be used in a way that will cause retaliative behaviors without the supervision of a human. These automated escalation loops obscure the boundaries between defense and offensive aggression and leave the possibility of unintended war among organizations, even nations.

The other fundamental problem is that of accountability and attribution. Existing legislation is based on recognizable operators of

cyber operations. As APTs and autonomous defenders engage without human control, there is an even higher challenge in pinpointing blame over breach, disruption or collateral damage. The law needs to change to identify liability where an autonomous agent makes harmful or erroneous decisions- whether the design, deployment or the organization that governs the environment.

The questions of privacy and data governance are also present. Defensive agents need to have access to large volumes of telemetry data such as user behavior logs, system metadata, network flows, and in some cases sensitive personal information to train and perform well. The privacy-protecting mechanisms would be insufficient to ensure that such data collection does not breach the compliance requirements provided in GDPR, HIPAA, and other data protection laws around the world. Also, vast amounts of telemetry bring an additional attack surface, posing a new threat to data security in case attackers can bypass protective systems. One of the major ethical issues is that of dual-use vulnerabilities. Adversaries can use defensive AI, particularly that that can identify vulnerabilities or create patches, to discover new vulnerabilities and use them. Such dual sense needs to be governed, controlled in terms of access, and model use limitations to ensure abuse.

In addition, systemic vulnerability can occur due to bias in AI-driven cyber defense. When the agents have been trained on imbalanced or biased data, they might still be more focused on specific alerts and overlook other alerts- leaving some infrastructure blind spots. It is vital to ensure fairness and minimize model drift as well as continuously retrain the systems with diverse and representative data.

Lastly, there should be international mechanisms to control independent cyber weapons, as there is on the case of kinetic warfare conventions. The development of offensive autonomous systems should be controlled by international treaties, ethical control committees, and the operation policies of AI in order to make sure that defensive agents act within the limits of legal regulations.

Discussion and Future Directions

The paradigm of Agent-vs- Agent Cyber Warfare is a radical change in offensive and defensive cybersecurity strategy. The growing autonomy of AI-powered APTs compels the defenders to go beyond the reactive control paradigm and take a more proactive, predictive, and self-defending defense architecture approach. A single key direction of future research consists of the creation of interoperable agent ecosystems capable of interoperating across distributed infrastructures of cloud, on-premises, IoT, and edge. As workloads move between hybrid platforms, defensive agents have to be able to keep situational awareness on a wide array of digital surfaces and keep pace with the changing configurations.

The second significant trend is the improvement of explainable agent behavior. The existing RL-driven or LLM driven agents tend to be black boxes, which creates uncertainties in transparency in decision-making. Future studies should develop more hybrid models that strike the right balance between autonomy and interpretability to allow analysts to audit and justify automated intervention, as well as verify intent. This is particularly important in the regulated industries like the healthcare sector, financial sector and military.

Also, bio-inspired agent intelligence will be useful in future cyber-defense systems, and the concepts will be based on immune systems, swarm behavior, and natural adaptive cycles. Independent agents that emulate biological ecosystems, such as self-healing, self-replication, and self-organization, are capable of offering a very resilient defense infrastructure to absorb and adapt to disruptive attack deployments by AI.

Quantum-secure agent design is also becoming popular. As quantum computing becomes a reality, attack agents can in the near future crack classical cryptography or find vulnerabilities at faster rates than ever before. Protective agents should thus incorporate PQC (Post-Quantum Cryptography) as well as quantum-resistant communication networks and quantum-driven detection software in order to be viable in the future threat environment.

Lastly, collaborative intelligence, i.e., a collaboration between human analysts and AI systems will be a crucial component. Humans are associated with contextual judgment, moral reasoning, and strategic supervision whereas agents have high-speed identification, high-scale analytics, and automatic execution. The combination of both of these capabilities will create a hybrid model of operation where the complex and the fast are handled by AI, and the intent and growth are controlled by people.

Conclusion

The advent of Agent-vs-Agent Cyber Warfare is a key turning point in the cybersecurity environment. With the growing autonomy, flexibility and persistence of AI-powered APTs, traditional security means based on manual analysis and fixed-point detection cannot be

relied upon any longer. The defensive systems should become fully autonomous agents with the ability to perceive any threat, make decisions in machine speed and learn through continuous confrontation behavior. The self-evolving cyber ecosystem that develops between offensive and defensive AI agents is a dynamic interaction of both parties, which develops in parallel.

The implementation of multi-agent defensive systems, which are supported by the optimization based on RL and LLLM enhanced reasoning and automated deception technologies, offers a plausible way to achieve resilient cybersecurity in an environment that is too complicated to be managed only by humans. Such agents are able to be proactive in finding out the vulnerabilities, predicting the strategies that the adversary might use, and implementing countermeasures before harm is inflicted. The increasing freedom of both attackers and defenders, however, also present their own major challenges- both legal grey areas and ethical questions to potential automated escalation and an unintended backlash.

Finally, the next stage of cybersecurity will not be determined by the use of better tools or more sophisticated algorithms, but rather by the hand of autonomous agents and the experience of a human being. The most promising way of preventing the emerging AI-driven APTs can be considered a balanced ecosystem, where scale, speed, and complexity are managed by AI and human judgment, oversight, and ethical governance. Companies that are preparing to enter this new era will need to invest in autonomous agents of defense, explainable AI, adversarial training, and ethical governance systems. Cyber warfare is becoming machine speed and only equally intelligent, adaptive and autonomous systems would be able to defend the digital infrastructures in the decades to come.

References

M. Wooldridge, *An Introduction to MultiAgent Systems*, 2nd ed., Wiley, 2021.

R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 2018.

N. Shaukat et al., "A Survey on Cybersecurity Challenges in Smart Grids," *Computers & Security*, 2020.

K. Kim et al., "AI-Based Cyber Attack Detection: A Survey," *IEEE Access*, 2021.

S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed., Pearson, 2021.

A. Martinez et al., "Deep Learning for Malware Detection," *IEEE TIFS*, 2020.

P. Brundage et al., "The Malicious Use of AI," OpenAI/CSER, 2018.

B. Biggio and F. Roli, "Wild Patterns: Adversarial Attacks," *Pattern Recognition*, 2018.

L. Yang et al., "Autonomous Cyber Defense using RL," *IEEE S&P Workshops*, 2020.

A. Chowdhary et al., "AI for Cyber Defense Systems," *Computers & Security*, 2021.

S. Kumar et al., "AI-Enabled APT Modeling," *Future Generation Computer Systems*, 2022.

S. Bhatia et al., "AI-Powered Spearphishing," *arXiv:1908.*, 2019.

S. E. McGregor, "AI and National Cyber Security," *Journal of Cyber Policy*, 2020.

M. Rigaki and S. Garcia, "AI-Guided Malware," *ECML PKDD*, 2018.

Google Brain, "RL for Autonomous Systems," 2020.

DARPA, "Cyber Grand Challenge Final Report," 2019.

MIT Lincoln Lab, "AI-Driven Cyber Ranges," Technical Report, 2021.

Y. Hu et al., "LLM-Based Phishing," *IEEE Access*, 2022.

L. Chen et al., "AI in Cloud Security," *IEEE Cloud Computing*, 2020.

R. Sommer et al., "Machine Learning in Cybersecurity," *USENIX*, 2019.

IBM X-Force, "AI-Powered Threat Intelligence," 2021.

C. Rebuffi et al., "Adversarial Robustness," *NeurIPS*, 2018.

A. Koroniotis et al., "Cyber-Kill Chain ML Modeling," *Computers & Security*, 2019.

T. S. Kumar et al., "Autonomous Response Agents," *IEEE Access*, 2021.

C. Feng et al., "Deep Learning IDS," *IEEE TII*, 2019.

C. S. Leung, "AI in Intrusion Detection," *Information Sciences*, 2020.

U.S. DoD, "AI-Driven Cyber Operations Report," 2021.

Z. Pawlicki et al., "Game-Theoretic Cyber Defense," *IEEE Transactions on Cybernetics*, 2019.

J. Camacho et al., "Autonomous Agents for Network Defense," *Sensors*, 2020.

M. Henry et al., "AI-Driven Deception Systems," *IEEE S&P Workshops*, 2021.

Check Point Research, "AI-Based Malware Trends," 2022.

McAfee Labs, "APTs and AI Automation," 2019.

CrowdStrike, "Adversary Tradecraft Report," 2021.

S. Gupta et al., "AI-Driven Threat Hunting," *IEEE Access*, 2020.

ENISA, "AI Threat Landscape," 2021.

CISA, "AI and Next-Gen Cyber Threats," 2022.

Vadisetty, R., Polamarasetti, A., Varadarajan, V., Kalla, D., Ramanathan, G.K. (2026). Cyber Warfare and AI Agents: Strengthening National Security Against Advanced Persistent Threats (APTs). In: Dhoska, K., Spaho, E. (eds) AI and Digital Transformation: Opportunities, Challenges, and Emerging Threats in Technology, Business, and Security. ICITTBT 2025. Communications in Computer and Information Science, vol 2669. Springer, Cham. https://doi.org/10.1007/978-3-032-07373-0_43

W. Wang et al., "Graph Neural Networks for Cybersecurity," *IEEE TNNLS*, 2021.

H. Zhang et al., "Adversarial ML Challenges," *Proceedings of IEEE*, 2019.

A. Alshamrani, "APT Attack Lifecycle Analysis," *Computers & Security*, 2021.

S. R. Rajbahadur et al., "RL for Security Automation," *IEEE Access*, 2022.