International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct-2026, PP 01-06

Artificial Intelligence Against Cyber Threats: A Review Of Detecting And Preventing Hacker Attacks Across Network, Endpoint And User Layers

Panchal Aditya Krishnakant[1] Prof. J. M. Patade[2] Prof. M. C. Kolse[3] Prof. S. K. Said [4]
Department of Artificial Intelligence and Data Science
Jaihind College of Engineering
Pune, Maharashtra, India

Abstract—Cyber threats and hacker attacks are increasing rapidly with the growth of internet use, online transactions, cloud computing and digital communication, and traditional security systems that depend on predefined rules and signatures often fail against advanced attacks. Artificial Intelligence offers a way to identify suspicious activities, detect unusual patterns and prevent attacks. This paper reviews AI for detecting and preventing hacker attacks across the network, endpoint and user layers. It examines machine learning for network intrusion detection, deep learning for malware and ransomware detection from executable files and file-system behaviour, anomaly detection on system logs, phishing detection and user behaviour analytics for unauthorized access. The effectiveness and limitations of AI in security operations are discussed, including evaluation pitfalls, false positives and adversarial attacks in which attackers manipulate inputs to evade detectors. Mapping of detections to attacker tactics and automated response are also covered. A five-layer AI cyber-defence architecture comprising data collection, AI detection engines, correlation and threat intelligence, automated response and human analysts is proposed, along with research gaps, objectives, a comparison of methods and challenges. The review concludes that AI must be integrated with traditional controls, human expertise and sound policy.

Index Terms—Cybersecurity, artificial intelligence, intrusion detection, malware detection, ransomware, anomaly detection, adversarial machine learning, SOC

I. Introduction

Attackers today combine phishing to steal credentials, malware to gain control of devices, ransomware to extort organizations and lateral movement to reach valuable data. Each stage leaves traces in network traffic, files, logs and user behaviour, but the volume of data in a modern organization is far too large for analysts to inspect manually.

AI helps by learning patterns of normal and malicious activity and highlighting what deserves attention. Buczak and Guven [1] surveyed machine learning for intrusion detection, and Apruzzese *et al.* [2] critically evaluated machine and deep learning across cybersecurity tasks, showing both strong results and significant limitations in practice.

This paper reviews AI techniques for detecting and preventing hacker attacks across layers and proposes an integrated defence architecture.

II. Related Work

Network. Sharafaldin *et al.* [3] provided CICIDS2017 for training intrusion detectors, and Mirsky *et al.* [4] proposed Kitsune for online unsupervised detection. Sommer and Paxson [5] cautioned about the difficulty of anomaly detection in real networks.

Endpoint. Saxe and Berlin [6] detected malware from static binary features with deep networks, and Raff *et al.* [7] trained a network on raw executable bytes. Anderson and Roth [8] released the EMBER dataset for static malware detection. For

ransomware, Kharraz *et al.* [9] analysed families and Continella *et al.* [10] built ShieldFS, which detects ransomware from file-system activity and restores affected files.

Logs and users. Du *et al.* [11] proposed DeepLog, an LSTM that learns normal log sequences and detects deviations indicating attacks. Isolation Forest [12] is widely used for user and entity behaviour anomalies.

Adversarial AI. Kolosnjaji *et al.* [13] showed that small byte changes can make malware evade deep detectors, and Biggio and Roli [14] reviewed a decade of adversarial machine learning. Table I summarizes these works.

Ref.	Author(s) / Year	Technique / Approach	Key Contribution	Limitation / Gap
[1]	Buczak and Guven, 2016	Survey of ML for intrusion detection	Mapped methods and datasets	Dataset-dependent results
[6]	Saxe and Berlin, 2015	Deep network on binary features	Accurate static malware detection	Evasion possible
[7]	Raff et al., 2018	MalConv on raw bytes	Feature-free malware detection	Adversarial vulnerability
[10]	Continella et al., 2016	ShieldFS ransomware defence	Detection and file recovery	Filesystem-level only
[11]	Du et al., 2017	DeepLog LSTM on logs	Detects abnormal log sequences	Log format dependence
[4]	Mirsky et al., 2018	Kitsune network IDS	Online unsupervised detection	Training contamination
[2]	Apruzzese et al., 2018	Evaluation of ML in security	Exposed practical limits	No new detector
[13]	Kolosnjaji et al., 2018	Adversarial malware binaries	Evaded deep malware detector	Attack demonstration

TABLE I Summary of Literature Survey

III. Preliminaries

3.1 The Attack Lifecycle

Fig. 1 maps attack stages, from reconnaissance and initial access to execution, persistence, lateral movement and impact, to the AI techniques that can detect each. MITRE ATT&CK [15] catalogues these tactics and techniques.

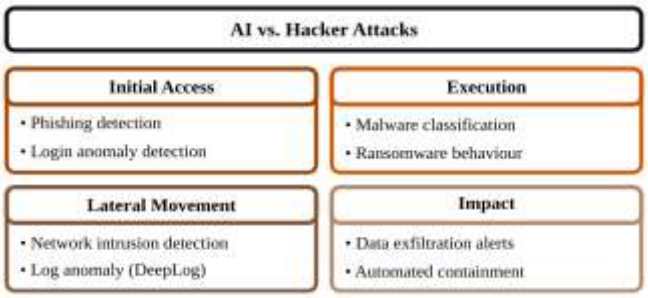


Fig. 1. Attack stages and AI detection techniques.

3.2 Detection Paradigms

Signature-based detection matches known indicators; anomaly-based detection flags deviations from learned normal behaviour; and supervised learning classifies activity based on labelled examples.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct-2026, PP 01-06

IV. Problem Statement and Objectives

4.1 Identified Research Gaps

- Models trained on benchmark datasets often fail in real environments.
- False positives overwhelm analysts and erode trust in AI alerts.
- Attackers adapt and craft adversarial inputs to evade detectors.
- Detections across network, endpoint and user layers are poorly correlated.
- Attackers increasingly use AI themselves, for example to write phishing.

4.2 Objectives of the Review

- To review AI methods for detecting attacks at the network, endpoint and user layers.
- To examine malware, ransomware and log-based detection.
- To analyse the limitations of AI in security, including adversarial attacks.
- To propose an integrated AI cyber-defence architecture.
- To identify challenges and future directions.

4.3 Proposed Framework

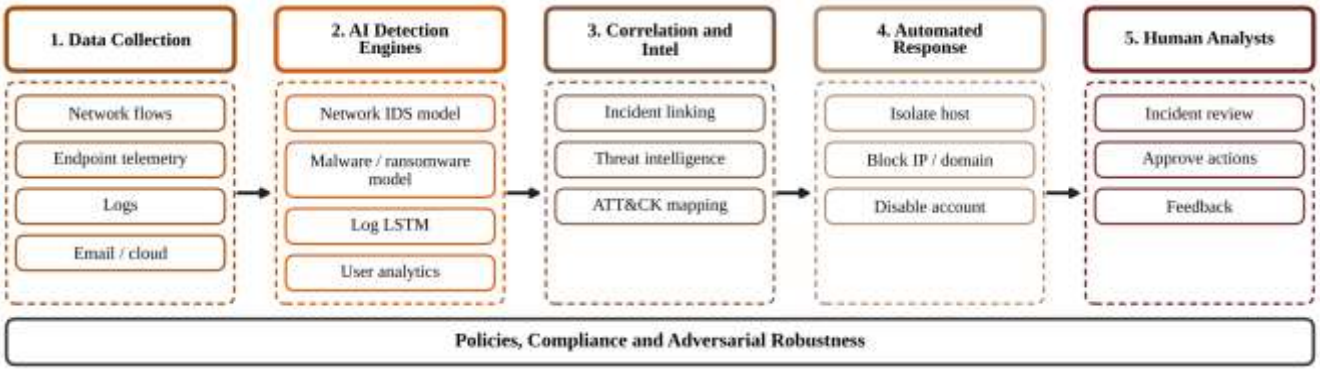


Fig. 2 shows the proposed architecture.

Fig. 2. Proposed integrated AI cyber-defence architecture.

4.4 Data Collection Layer

Network flows, endpoint telemetry, system and authentication logs, email and cloud activity are collected centrally.

4.5 AI Detection Engines Layer

A network model detects intrusions, an endpoint model classifies malware and ransomware behaviour, a log model detects abnormal sequences and a user analytics model flags unusual logins and access.

4.6 Correlation and Threat Intelligence Layer

Alerts are linked across layers into incidents, enriched with threat intelligence and mapped to ATT&CK techniques.

4.7 Automated Response Layer

Playbooks isolate infected hosts, block malicious IPs and domains, disable compromised accounts and restore files from backups.

4.8 Human Analyst Layer

Analysts review incidents, approve high-impact actions and provide feedback that improves the models.

V. Working Methodology

Fig. 3 describes the workflow. In DeepLog-style detection, an LSTM predicts the probability of the next log event; an event is anomalous if it is not among the top-g predicted events,

$$anomaly\ if\ e_t \notin Top_g(\ P(e \mid e_{t-h} \ldots e_{t-1})\ ) \qquad (1)$$

and incident risk combines alert scores across layers,

$$R = 1 - \Pi_i\ (1 - p_i) \qquad (2)$$

where p_i is the confidence of alert i belonging to the incident.

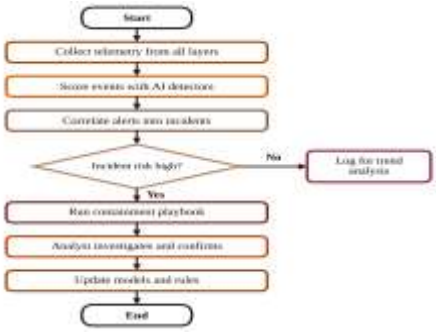


Fig. 3. AI-driven threat detection and response workflow.

VI. Discussion

6.1 Comparison of Approaches

Table II compares AI methods by layer. Network models detect scanning and intrusion, endpoint models detect malware and ransomware, log models catch abnormal system behaviour, and user analytics detect account compromise.

Layer	AI Method	Detects	Limitation
Network	RF / autoencoder IDS	Scans, intrusions, DDoS	Encrypted traffic
Endpoint	Deep malware classifiers	Malware, ransomware	Adversarial evasion
Logs	LSTM sequence models	Abnormal behaviour	Log changes
Users	Behaviour analytics	Account compromise	Insider variability
Email	NLP classifiers	Phishing	AI-written lures

TABLE II Comparison of AI Methods by Defence Layer

6.2 Open Challenges and Future Work

Table III summarizes challenges. Future directions include adversarially robust models, AI assistants for analysts, federated threat intelligence and defences against AI-generated attacks.

Challenge	Mitigation
False positives	Cross-layer correlation; tuning
Generalization	Local retraining; realistic data
Adversarial evasion	Robust training; model diversity
Analyst overload	Prioritization; AI assistants
AI-powered attackers	Continuous threat intelligence

TABLE III Challenges and Mitigation

VII. Conclusion and Future Work

AI is now essential for defending against cyber threats at the scale and speed of modern attacks. The reviewed literature shows effective AI detection of intrusions, malware, ransomware, log anomalies and account compromise, but also false positives, generalization gaps and adversarial evasion. An integrated architecture that correlates AI detections, automates safe responses and keeps human analysts in control offers reliable protection.

7.1 Acknowledgment

The authors express their sincere gratitude to the Department of Artificial Intelligence and Data Science, Jaihind College of Engineering, Pune, for providing the facilities, guidance and encouragement needed to carry out this work. The authors also thank the faculty members and colleagues whose valuable suggestions helped improve the quality of this review.

References

1. A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” IEEE Commun. Surveys Tuts., vol. 18, no. 2, pp. 1153–1176, 2016.
2. G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, “On the effectiveness of machine and deep learning for cyber security,” in Proc. 10th Int. Conf. Cyber Conflict (CyCon), 2018, pp. 371–390.
3. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in Proc. 4th Int. Conf. Inf. Syst. Security Privacy (ICISSP), 2018, pp. 108–116.
4. Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, “Kitsune: An ensemble of autoencoders for online network intrusion detection,” in Proc. Netw. Distrib. Syst. Security Symp. (NDSS), 2018.
5. R. Sommer and V. Paxson, “Outside the closed world: On using machine learning for network intrusion detection,” in Proc. IEEE Symp. Security Privacy, 2010, pp. 305–316.
6. J. Saxe and K. Berlin, “Deep neural network based malware detection using two dimensional binary program features,” in Proc. 10th Int. Conf. Malicious Unwanted Softw. (MALWARE), 2015, pp. 11–20.
7. E. Raff et al., “Malware detection by eating a whole EXE,” in Proc. AAAI Workshops, 2018.
8. H. S. Anderson and P. Roth, “EMBER: An open dataset for training static PE malware machine learning models,” arXiv:1804.04637, 2018.
9. A. Kharraz, W. Robertson, D. Balzarotti, L. Bilge, and E. Kirda, “Cutting the Gordian knot: A look under the hood of ransomware attacks,” in Proc. DIMVA, 2015, pp. 3–24.
10. A. Continella et al., “ShieldFS: A self-healing, ransomware-aware filesystem,” in Proc. 32nd Annu. Comput. Security Appl. Conf. (ACSAC), 2016, pp. 336–347.

11. M. Du, F. Li, G. Zheng, and V. Srikumar, “DeepLog: Anomaly detection and diagnosis from system logs through deep learning,” in Proc. ACM CCS, 2017, pp. 1285–1298.

12. F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation forest,” in Proc. 8th IEEE Int. Conf. Data Min. (ICDM), 2008, pp. 413–422.

13. B. Kolosnjaji et al., “Adversarial malware binaries: Evading deep learning for malware detection in executables,” in Proc. 26th Eur. Signal Process. Conf. (EUSIPCO), 2018, pp. 533–537.

14. B. Biggio and F. Roli, “Wild patterns: Ten years after the rise of adversarial machine learning,” Pattern Recognit., vol. 84, pp. 317–331, 2018.

15. B. E. Strom et al., “MITRE ATT&CK: Design and philosophy,” MITRE Corp., Tech. Rep., 2018.