International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/

Vol. 58 No. 2 August 2026, pp. 983-996

From Intrusion Classification to Threat Forecasting: A Systematic Mapping Review of Machine Learning Research

Alih Rama Hamda[1], Bisyron Wahyudi[2]

[1,2]Cyber Defense Engineering, Republic of Indonesia Defense University, Indoneisa

Auteur correspondant : Alih Rama Hamda, arhmda4@gmail.com

Abstract: This review examines whether machine-learning intrusion detection studies genuinely forecast future cyber threats or mainly classify current observations, and evaluates their readiness for cyber threat intelligence operations. A systematic mapping review used two Scopus searches covering English articles and reviews from 2015 to 2025. The searches returned 310 records. After seven duplicates were removed, 303 records were screened. Twenty-two secondary studies were separated, and 281 primary studies were coded by publication characteristics, methods, datasets, evaluation evidence, temporal objective, and operational readiness. Publications increased from two in 2016 to seventy-seven in 2025. Of the primary studies, 260 performed contemporaneous detection or classification, eleven forecast future attack sequences, plans, targets, or trends, seven addressed early or pre-event detection, and three implemented adaptive or continual learning. Accuracy was reported far more often than latency, external validation, statistical uncertainty, or integration with security operations. Explicit cyber threat intelligence terminology appeared in eleven studies, while direct integration with security information and event management or security operations centers was rare. The literature is expanding rapidly but remains dominated by benchmark-based classification rather than operational forecasting. This review contributes a task-timing taxonomy and a five-dimensional readiness framework covering temporal validity, data realism, deployment feasibility, analyst interpretability, and cyber threat intelligence integration. Future studies should report prediction horizon, lead time, chronological validation, calibration, latency, and operational utility.

Keywords: cyber threat intelligence; early warning; concept drift; benchmark dataset; operational readiness; attack sequence; explainable artificial intelligence.

1. INTRODUCTION

Machine-learning intrusion detection has grown from a specialist research topic into a major component of network, cloud, Internet of Things (IoT), industrial, and cyber-physical security. Recent reviews document broad progress in supervised, unsupervised, deep, adversarial, and explainable intrusion-detection methods, but they also report persistent dependence on benchmark datasets and heterogeneous evaluation protocols [1], [2], [3], [4]. Cyber threat intelligence extends the objective beyond identifying a malicious flow: it seeks to anticipate adversary behavior, contextualize indicators, support prioritization, and inform defensive action [5], [6].

A central conceptual problem remains unresolved. In machine learning, a classifier “predicts” a label for the current observation; in security operations, prediction usually implies a future-oriented judgment, such as the next attack step, a likely target, an attack window, or an actionable early warning. Treating these meanings as equivalent inflates claims of predictive capability. The distinction matters because operational CTI requires lead time, temporal validation, evolving-threat adaptation, and integration with analyst workflows, not accuracy alone. The alert-correlation literature has recognized the importance of future attack prediction, yet broad IDS reviews still tend to organize evidence by algorithm, application domain, or dataset [7], [8], [9].

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 2 August 2026, pp. 983-996

This study therefore reframes the literature through both a task-oriented and operation-oriented lens. Rather than assuming that every classification model is predictive, it asks whether the output is contemporaneous, early, adaptive, or explicitly future-oriented. It then examines whether studies report evidence needed for practical CTI deployment, including latency, throughput, cross-dataset validity, explainability, adversarial robustness, and integration with CTI, SIEM, SOC, or MITRE ATT&CK concepts.

The contributions are fourfold. First, the study provides a dual-query Scopus corpus that covers predictive IDS and explicit CTI contexts. Second, it introduces a conservative temporal taxonomy that prevents "prediction" from being used as a synonym for classification. Third, it quantifies the gap between benchmark performance and operational readiness. Fourth, it compares the resulting scope with prior reviews and derives a focused research agenda for genuine cyber-attack forecasting and early warning.

1.1. Research Questions

Table 1: Research questions

RQ	Question
RQ1	How did predictive IDS and CTI-related research evolve from 2015 to 2025?
RQ2	Which algorithms, datasets, application domains, and evaluation practices dominate the corpus?
RQ3	How many studies implement genuine temporal forecasting, early/pre-event detection, or adaptation rather than contemporaneous classification?
RQ4	To what extent do the studies provide evidence of operational CTI readiness?
RQ5	How does this review differ from recent IDS, CTI, explainable-AI, dataset, and attack-prediction reviews?

2. MATERIAL AND METHOD

2.1. Review Design and Scope

The study was designed as a systematic mapping review, following the transparency principles of PRISMA 2020 and software-engineering SLR guidance [10], [11]. A mapping review was selected because the objective was to characterize a heterogeneous body of predictive, classification, early-warning, and adaptive IDS research rather than pool commensurable effect sizes. The unit of analysis was a Scopus-indexed publication whose title, abstract, or keywords addressed cyber or network intrusion detection together with machine learning, deep learning, statistical learning, or predictive analytics.

2.2. Search Strategy

The Scopus search was executed on 28 July 2026 and restricted to English-language articles and reviews published from 2015 through 2025. Search S1 targeted predictive network or cyber intrusion detection. Search S2 targeted explicit CTI, SIEM, SOC, or threat-intelligence contexts. The two queries were intentionally separated: requiring CTI terminology in the main query would have reduced recall because many future-oriented IDS studies do not label themselves as CTI. A third search (S3) collected prior reviews published from 2020 to 2025 for the novelty analysis; S3 was not merged into the primary corpus.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 2 August 2026, pp. 983-996

The full Boolean strings and raw Scopus CSV exports were retained for reproducibility. S1 and S2 produced 310 records. Deduplication used DOI, Scopus EID, and normalized title, removing seven overlaps and leaving 303 unique records.

Table 2: Scopus search strategy and record counts

Search	Purpose	Condensed search logic	Records
S1	Predictive IDS	Network/cyber intrusion detection AND predictive/early/forecasting terms AND ML/DL/statistical learning; 2015–2025; article/review; English	290
S2	Operational CTI	CTI/threat intelligence/SOC/SIEM AND intrusion/threat detection AND predictive/early terms AND AI/ML; same limits	20
S3	Prior reviews	IDS/cyber-threat detection AND review/survey/bibliometric terms AND ML/DL/predictive terms; 2020–2025	239

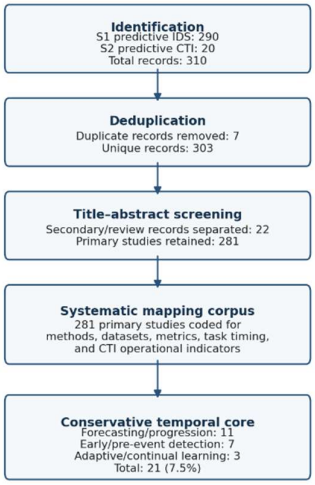

Figure 1: Identification, screening, and temporal-core selection

2.3. Eligibility and Coding

Records were first separated into primary and secondary evidence. Seventeen records were indexed by Scopus as reviews. Five additional records indexed as articles were explicitly titled as a review, survey, overview, taxonomy, or roadmap. These 22 secondary studies were excluded from the primary mapping corpus but retained for contextual comparison, leaving 281 primary studies.

A conservative coding rule was applied: the words predict, prediction, or predictive did not by themselves establish temporal prediction. A paper was coded as forecasting only when its title or abstract stated a future-oriented target or mechanism. Ambiguous cases were placed in contemporaneous detection to avoid overstating predictive capability. The coding therefore represents evidence visible in Scopus titles, abstracts, and keywords, not an inference from algorithm names.

Table 3: Eligibility and temporal-coding rules

Rule	Operational definition
Include	Cyber/network intrusion, malicious-traffic, attack, or security-event detection; ML/DL/statistical method; primary article; 2015–2025; English.
Exclude from primary corpus	Review/survey/overview/taxonomy; non-cyber intrusion; purely cryptographic, authentication, or policy study without intrusion/attack analytics.
Temporal forecasting/progression	Explicit prediction of a future attack, next target, next stage, attack plan, alert sequence, attacker behavior, or long-horizon threat trend.
Early/pre-event detection	Detection before session completion, before encryption/damage, during an initial attack phase, or with an explicit early-warning objective.
Adaptive/continual learning	Concept-drift handling, class-incremental learning, continual learning, or model/signature updating from newly observed attacks.
Contemporaneous detection/classification	Classification or anomaly detection of the current flow, packet, session, or event without a clear future, early, or adaptive target.

2.4. Data Extraction and Analysis

The extraction matrix included year, source, citations, affiliations, title, abstract, author and index keywords, document type, DOI, query source, temporal category, algorithm mentions, dataset mentions, evaluation metrics, application domain, and operational indicators. Publication trends and frequencies were calculated by whole counting at the document level. Country counts used whole counting of unique countries listed in each record's affiliations, so totals can exceed the number of studies. Keyword and phrase matching was normalized for common abbreviations such as SVM, CNN, LSTM, CTI, SIEM, and XAI. Co-word mapping principles were informed by established bibliometric practice [12], while all reported frequencies were recalculated directly from the Scopus metadata.

Operational readiness was assessed through five dimensions: (1) temporal validity future horizon, early lead time, or adaptation; (2) data realism recent, multi-source, cross-dataset, or externally validated data; (3) deployment feasibility real-time operation, latency, throughput, resource cost, or production evidence; (4) analyst interpretability explainability, calibration, and decision-relevant outputs; and (5) CTI integration explicit threat intelligence, ATT&CK/kill-chain mapping, SIEM, SOC, or response orchestration. Because the studies used heterogeneous datasets and protocols, no meta-analysis of accuracy or F1-score was attempted.

3. RESULT

3.1. Corpus Growth, Sources, and Geography

The final primary corpus contained 281 studies. Annual output rose from 2 studies in 2016 to 77 in 2025, with the steepest sustained growth beginning in 2022. The final three years accounted for 184 studies (65.5% of the corpus), indicating that predictive language and AI-based IDS research accelerated substantially during 2023–2025.

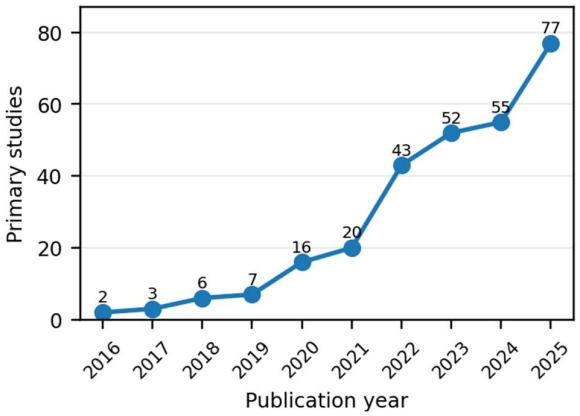

Figure 2: Annual distribution of the 281 primary studies

Table 4: Most frequent publication sources

Source	Studies
IEEE Access	17
Electronics (Switzerland)	10
Applied Sciences (Switzerland)	9
International Journal of Advanced Computer Science and Applications	7
Computers and Security	6
Computers, Materials and Continua	6
Sensors	5
Multimedia Tools and Applications	5
Scientific Reports	5
Cluster Computing	4

IEEE Access was the most frequent source, followed by Electronics, Applied Sciences, and several security- and computing-oriented journals. Affiliation whole counting showed the largest contributions from India (96 records), Saudi Arabia (37), China (28), the United States (24), and the United Kingdom (19). These counts describe participation, not exclusive national ownership, because internationally co-authored papers contribute to multiple countries.

Table 5: Leading affiliation countries using whole counting

Country	Studies
India	96
Saudi Arabia	37
China	28
United States	24
United Kingdom	19
South Korea	14
Pakistan	13
Canada	10
Iran	9
Turkey	9

3.2. Methods, Datasets, and Evaluation Practices

Method mentions were highly diverse, but the most common families remained conventional supervised learners and deep sequential models. Random forest appeared in 40 studies, decision trees in 37, CNNs and ensemble methods in 33 each, SVMs and LSTMs in 32 each, and deep neural networks in 28. Newer approache federated learning, transformers/LLMs, graph neural networks, and reinforcement learning were visible but much less frequent.

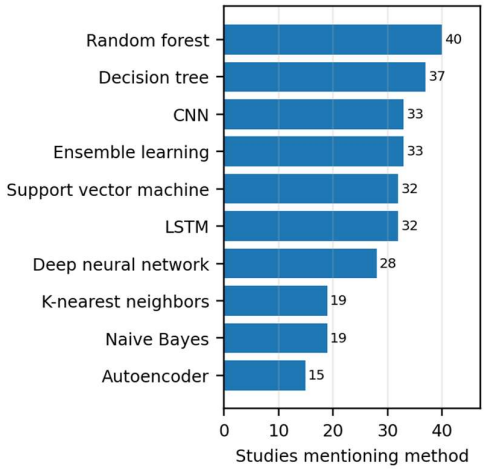

Figure 3: Ten most frequently mentioned method families

Benchmark concentration remained pronounced. NSL-KDD and UNSW-NB15 were each mentioned by 29 studies, followed by CICIDS2017 (14), CIC-DDoS2019 (10), CSE-CIC-IDS2018 (7), TON_IoT (7), and BoT-IoT (5). This concentration supports the

concern raised by recent dataset reviews that benchmark familiarity can improve comparability while limiting ecological validity [8].

Table 6: Evaluation evidence visible in titles, abstracts, and keywords

Evaluation item	Studies
Accuracy	184
Precision	62
Recall	48
F1-score	44
False-positive rate	14
ROC-AUC	11
Latency / inference time	11
Throughput	8
Cross-validation	10
Statistical significance / CI	2

Accuracy was mentioned by 184 studies, far exceeding precision (62), recall (48), F1-score (44), false-positive rate (14), and ROC-AUC (11). Only 11 studies mentioned latency or inference time, eight throughput, ten cross-validation, and two statistical significance or confidence intervals. The imbalance indicates that predictive IDS evaluation remains dominated by point performance measures rather than uncertainty, temporal utility, and deployment cost.

3.3. Classification Is Not the Same as Threat Forecasting

The temporal audit produced the clearest result of the review. Of 281 primary studies, 260 (92.5%) classified the current packet, flow, session, or event without a clear future, early, or adaptive target. Only 11 (3.9%) explicitly forecast a future attack sequence, plan, target, attacker preference, alert, defensive agility, or longer-term trend. Seven (2.5%) implemented early or pre-event detection, and three (1.1%) implemented adaptive or continual learning.

Table 7: Conservative task-timing taxonomy

Task category	Studies	Share
Contemporaneous detection/classification	260	92.5%
Temporal forecasting/progression	11	3.9%
Early/pre-event detection	7	2.5%
Adaptive/continual learning	3	1.1%

The 11 future-oriented studies illustrate what genuine prediction looks like. DeepOP predicts MITRE ATT&CK sequences using deep learning and ontology [13]. Attack-graph and attack-plan studies model exploit paths and upcoming phases [14], [15]. GRU and hidden-Markov approaches forecast future alerts or attack stages from historical sequences [16], [17], [18]. Other work predicts long-term cyber-threat trends, defense agility, attacker preferences, future institutional targets, or context-dependent attacks [19], [20], [21], [22], [23].

Early-warning studies used a different operational objective: detect an attack before the complete session, damaging phase, or full attack life cycle. Examples include proactive SDN warning, digital-forensics detection of an ongoing attack, early cyber-physical attack detection, packet-level GAN classification before session completion, healthcare edge detection, pre-encryption ransomware detection, and adjustable early graph-based NIDS [24], [25], [26], [27], [28], [29], [30]. Adaptive studies addressed class-incremental learning, concept drift, or automatic signature/model updating [31], [32], [33].

This distinction also exposes a terminology problem. Fifty-nine studies in the initial lexical coding used prediction-oriented language, yet only 11 stated a future attack, sequence, target, plan, behavior, or trend. Thus, at least 48 of the 59 prediction-labeled studies (81.4%) used prediction primarily in the statistical sense of assigning a current class label. This is not methodologically wrong, but it should not be presented as proactive threat forecasting.

3.4. Operational Cyber Threat Intelligence Readiness

Operational evidence was sparse relative to the volume of algorithmic work. Real-time or online operation was mentioned by 57 studies (20.3%), but latency or inference time appeared in only 11 (3.9%). Resource cost or throughput appeared in 25 (8.9%). Explainability or XAI appeared in 13 (4.6%), while cross-dataset or external validation appeared in four (1.4%). Explicit CTI or threat-intelligence terminology appeared in 11 studies (3.9%), MITRE ATT&CK or kill-chain concepts in five (1.8%), SIEM in two (0.7%), and SOC in one (0.4%).

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 2 August 2026, pp. 983-996

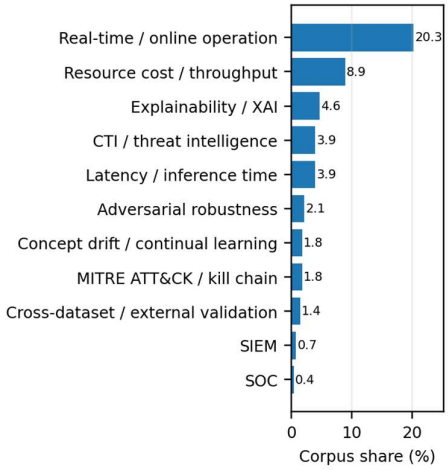

Figure 4: Operational-readiness indicators in the primary corpus

These results show that a model can achieve high offline accuracy without demonstrating operational value. For CTI use, a result must be timely, calibrated, attributable to a time horizon, robust to data drift, interpretable to analysts, and connected to a workflow that changes defensive decisions. The rarity of SIEM/SOC integration and external validation suggests that most studies remain at the model or benchmark layer rather than the intelligence-production and response layer.

3.5. Comparison with Prior Reviews and Novelty

The S3 search identified a large review landscape. Recent reviews cover general NIDS methods, meta-analysis, IoT deep learning, datasets, XAI, APT detection, CTI mining, and LLM-based threat detection [1], [2], [3], [4], [5], [6], [8], [9], [34]. The closest prior work is the alert-correlation survey, which specifically examines attack prediction [7]. The present review extends that line by auditing a broader ML-based IDS corpus and quantifying how often prediction claims actually contain a temporal target and operational CTI evidence.

Table 8: Positioning against representative prior reviews

Review	Primary focus	Temporal contribution	Operational/CTI emphasis
Abdulganiyu et al., 2023 [1]	General ML/DL NIDS SLR	No explicit task-timing audit	Limited
Maseer et al., 2024 [2]	Meta-analysis, datasets, validation	Evaluation-centric	Limited
Albasheer et al., 2022 [7]	Alert correlation and attack prediction	Strong on attack prediction	Partial
Sun et al., 2023 [5]	CTI mining and proactive defense	CTI-centric, not IDS corpus audit	Strong

Review	Primary focus	Temporal contribution	Operational/CTI emphasis
Neupane et al., 2022 [4]	Explainable IDS	No temporal taxonomy	Interpretability
Goldschmidt & Chudá, 2025 [8]	Intrusion datasets	No task/operations audit	Data realism
Salim et al., 2023 [6]	APT detection and situational awareness	APT-specific	Partial
This review	Task timing + CTI readiness across 281 studies	Quantifies classification–forecasting gap	Five-dimensional framework

4. DISCUSSION

4.1. What the Field Currently Predicts

The evidence indicates that the field mostly predicts a label, not a future security state. This explains why classification accuracy is abundant while lead time, horizon, calibration, and decision utility are rare. In conventional supervised IDS, the target is whether the current observation is benign or malicious. In temporal threat prediction, the target may be the next ATT&CK technique, the next affected asset, the probability of an attack within a time window, or the likely progression of a multi-stage campaign. Those tasks require different dataset construction, train–test splitting, leakage controls, and evaluation metrics.

A high-quality temporal evaluation should preserve chronology and report the prediction horizon, lead time, event prevalence, calibration, and utility of acting on the forecast. Random splits can leak future patterns into training and exaggerate performance. Similarly, accuracy can be misleading in highly imbalanced environments. Time-dependent precision–recall, Brier score, calibration error, expected cost, time-to-detection, and analyst workload are more relevant complements to standard classification metrics.

4.2. Benchmark Monoculture and External Validity

The dominance of NSL-KDD, UNSW-NB15, and CIC-family datasets supports reproducibility but creates a benchmark monoculture. Many datasets are static, laboratory-generated, or weakly connected to current enterprise architectures. Genuine CTI models need heterogeneous inputs: flows, endpoint events, identity telemetry, vulnerability information, asset criticality, threat reports, and historical campaigns. The data should be partitioned temporally and, where possible, validated across organizations or datasets. Only four studies signaled cross-dataset or external validation at abstract level, making generalization a major unresolved risk.

4.3. From Model Performance to Cyber Threat Intelligence Decision Support

Operational CTI is a socio-technical process. A useful model should not merely produce an attack class; it should provide evidence that an analyst can interpret and act upon. The proposed readiness framework is therefore organized around five dimensions.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 58 No. 2 August 2026, pp. 983-996

Table 9: Proposed predictive-intrusion-detection readiness framework for cyber threat intelligence operations

Dimension	Minimum evidence
Temporal validity	Clearly defined future/early/adaptive target; chronological split; horizon and lead-time reporting.
Data realism	Recent and heterogeneous telemetry; class prevalence; cross-dataset or external validation; drift monitoring.
Deployment feasibility	Latency, throughput, resource use, failure modes, and evidence from real-time or production operation.
Analyst interpretability	Calibrated confidence, explanation, attack context, uncertainty, and measurable effect on analyst workload.
CTI integration	Mapping to ATT&CK/kill chain; ingestion of threat intelligence; SIEM/SOC workflow; response or prioritization output.

The framework is intended as an evaluation checklist, not a new maturity standard. It helps distinguish a promising experimental classifier from a deployable predictive-intelligence capability. A study can be methodologically strong at the algorithm level while still lacking data realism or workflow evidence; the dimensions should therefore be reported separately rather than collapsed into a single score.

4.4. Research Agenda

Future research should prioritize four lines of work. First, attack prediction should be formulated as time-to-event, sequence, graph, or point-process modeling with explicit horizons and uncertainty. Second, external and longitudinal validation should replace single-dataset random splits, with concept-drift monitoring and safe online updating. Third, evaluation should include calibration, false-alert cost, latency, throughput, analyst workload, and response utility. Fourth, predictive outputs should be integrated with CTI and operational structures, including ATT&CK techniques, asset context, vulnerability information, SIEM enrichment, SOC triage, and human-in-the-loop explanation.

Emerging transformers, graph neural networks, federated learning, and LLMs may support multi-source reasoning, but model complexity should not substitute for temporal validity. Recent LLM and XAI surveys similarly emphasize interpretability and evaluation challenges [4], [34]. The decisive question is not whether a model is sophisticated, but whether it predicts the right future event early enough and reliably enough to improve a defensive decision.

4.5. Limitations and Validity Threats

This review has five limitations. First, Scopus was the only bibliographic database, so relevant IEEE, ACM, Web of Science, or gray-literature records may have been missed. Second, only English-language publications from 2015–2025 were included. Third, the systematic mapping relied on titles, abstracts, keywords, and bibliographic metadata; some full texts may contain temporal or operational details not visible in the abstract. The conservative rule therefore favors specificity over sensitivity. Fourth, keyword frequencies reflect mentions, not necessarily the main experimental contribution. Fifth, citation counts and 2025 publication totals are time-dependent. These limitations are partly mitigated by preserving the raw exports, using dual searches, documenting the coding rules, and avoiding quantitative pooling of heterogeneous performance results.

5. CONCLUSIONS

This systematic mapping review analyzed 281 primary studies on machine-learning intrusion detection and CTI-related predictive analytics published from 2015 to 2025. The literature expanded rapidly, but its predictive claim is much stronger in terminology than in temporal design. A conservative audit found that 92.5% of studies performed contemporaneous detection or classification, while only 3.9% forecast future attacks or progressions, 2.5% supported explicit early/pre-event detection, and 1.1% implemented adaptive or continual learning. Evaluation was dominated by accuracy and benchmark datasets, whereas latency, external validation, statistical uncertainty, explainability, and CTI/SIEM/SOC integration were uncommon.

The article's main contribution is a task-timing taxonomy and a five-dimensional CTI-readiness framework that separates statistical label prediction from operational threat forecasting. Research should move beyond isolated benchmark accuracy toward chronological validation, calibrated future-event prediction, cross-environment generalization, explainable decision support, and integration with threat-intelligence and security-operations workflows. Under this stricter definition, predictive IDS becomes not merely a more accurate detector, but an intelligence capability that creates reliable lead time for defensive action.

ACKNOWLEDGMENT

The authors acknowledge the Cyber Defense Engineering Program, Republic of Indonesia Defense University, for academic support. Bibliographic metadata were obtained from Scopus, and the authors retained the raw CSV exports and coding matrix to support reproducibility.

DECLARATIONS

AI USAGE STATEMENT

The authors declare that AI-assisted technologies (e.g., ChatGPT, OpenAI) were used to support the drafting and editing of this manuscript, specifically in refining grammar, improving sentence clarity, and checking coherence. No part of the conceptual framework, data interpretation, or conclusions was generated by AI. The authors have thoroughly reviewed and verified all content, and accept full responsibility for the work presented.

REFERENCES

[1] Abdulganiyu O.H., Ait Tchakoucht T., and Saheed Y.K., "A systematic literature review for network intrusion detection system (IDS)," International Journal of Information Security, vol. 22, no. 5, pp. 1125–1162, 2023. doi: 10.1007/s10207-023-00682-2.

[2] Maseer Z.K., Kadhim Q.K., Al-Bander B., Yusof R., and Saif A., "Meta-analysis and systematic review for anomaly network intrusion detection systems: Detection methods, dataset, validation methodology, and challenges," IET Networks, vol. 13, no. 5-6, pp. 339–376, 2024. doi: 10.1049/ntw2.12128.

[3] Chou D. and Jiang M., "A Survey on Data-driven Network Intrusion Detection," ACM Computing Surveys, vol. 54, no. 9, Art. no. 182, 2022. doi: 10.1145/3472753.

[4] Neupane S., Ables J., Anderson W., Mittal S., Rahimi S., Banicescu I., and Seale M., "Explainable Intrusion Detection Systems (X-IDS): A Survey of Current Methods, Challenges, and Opportunities," IEEE Access, vol. 10, pp. 112392–112415, 2022. doi: 10.1109/ACCESS.2022.3216617.

[5] Sun N., Ding M., Jiang J., Xu W., Mo X., Tai Y., and Zhang J., "Cyber Threat Intelligence Mining for Proactive Cybersecurity Defense: A Survey and New Perspectives," IEEE Communications Surveys and Tutorials, vol. 25, no. 3, pp. 1748–1774, 2023. doi: 10.1109/COMST.2023.3273282.

[6] Salim D.T., Singh M.M., and Keikhosrokiani P., "A systematic literature review for APT detection and Effective Cyber Situational Awareness (ECSA) conceptual model," Heliyon, vol. 9, no. 7, Art. no. e17156, 2023. doi: 10.1016/j.heliyon.2023.e17156.

[7] Albasheer H., Siraj M.M., Mubarakali A., Tayfour O.E., Salih S., Hamdan M., Khan S., Zainal A., and Kamarudeen S., "Cyber-Attack Prediction Based on Network Intrusion Detection Systems for Alert Correlation Techniques: A Survey," Sensors, vol. 22, no. 4, Art. no. 1494, 2022. doi: 10.3390/s22041494.

[8] Goldschmidt P. and Chudá D., "Network intrusion datasets: A survey, limitations, and recommendations," Computers and Security, vol. 156, Art. no. 104510, 2025. doi: 10.1016/j.cose.2025.104510.

[9] Alsoufi M.A., Razak S., Siraj M.M., Nafea I., Ghaleb F.A., Saeed F., and Nasser M., "Anomaly-based intrusion detection systems in iot using deep learning: A systematic literature review," Applied Sciences (Switzerland), vol. 11, no. 18, Art. no. 8383, 2021. doi: 10.3390/app11188383.

[10] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, Art. no. n71, 2021. doi: 10.1136/bmj.n71.

[11] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," EBSE Technical Report EBSE-2007-01, Keele University and Durham University, 2007.

[12] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," Scientometrics, vol. 84, no. 2, pp. 523–538, 2010. doi: 10.1007/s11192-009-0146-3.

[13] Zhang S., Xue X., and Su X., "DeepOP: A Hybrid Framework for MITRE ATT&CK Sequence Prediction via Deep Learning and Ontology," Electronics (Switzerland), vol. 14, no. 2, Art. no. 257, 2025. doi: 10.3390/electronics14020257.

[14] Mishra S., "Cyber defence using attack graphs prediction and visualisation," International Journal of Communication Networks and Distributed Systems, vol. 29, no. 3, pp. 268–289, 2023. doi: 10.1504/IJCNDS.2023.130566.

[15] Li T., Liu Y., Liu Y., Xiao Y., and Nguyen N.A., "Attack plan recognition using hidden Markov and probabilistic inference," Computers and Security, vol. 97, Art. no. 101974, 2020. doi: 10.1016/j.cose.2020.101974.

[16] Ansari M.S., Bartoš V., and Lee B., "GRU-based deep learning approach for network intrusion alert prediction," Future Generation Computer Systems, vol. 128, pp. 235–247, 2022. doi: 10.1016/j.future.2021.09.040.

[17] Zhang K., Zhao F., Luo S., Xin Y., Zhu H., and Chen Y., "Online intrusion scenario discovery and prediction based on hierarchical temporal memory (HTM)," Applied Sciences (Switzerland), vol. 10, no. 7, Art. no. 2596, 2020. doi: 10.3390/app10072596.

[18] Chadza T., Kyriakopoulos K.G., and Lambotharan S., "Learning to Learn Sequential Network Attacks Using Hidden Markov Models," IEEE Access, vol. 8, pp. 134480–134497, 2020. doi: 10.1109/ACCESS.2020.3011293.

[19] Patil G., Sapkal A., and Ingale V.S., "Designing an Improved Cyberattack Prediction Model Using Context-Aware Behavioral Modeling Analysis," Engineering, Technology and Applied Science Research, vol. 15, no. 5, pp. 27464–27469, 2025. doi: 10.48084/etasr.11799.

[20] Almahmoud Z., Yoo P.D., Alhussein O., Farhat I., and Damiani E., "A holistic and proactive approach to forecasting cyber threats," Scientific Reports, vol. 13, no. 1, Art. no. 8049, 2023. doi: 10.1038/s41598-023-35198-1.

[21] Muhati E. and Rawat D.B., "Hidden-Markov-Model-Enabled Prediction and Visualization of Cyber Agility in IoT Era," IEEE Internet of Things Journal, vol. 9, no. 12, pp. 9117–9127, 2022. doi: 10.1109/JIOT.2021.3056118.

[22] Veksler V.D., Buchler N., LaFleur C.G., Yu M.S., Lebiere C., and Gonzalez C., "Cognitive Models in Cybersecurity: Learning From Expert Analysts and Predicting Attacker Behavior," Frontiers in Psychology, vol. 11, Art. no. 1049, 2020. doi: 10.3389/fpsyg.2020.01049.

[23] Zabarah S., Naman O., Salahuddin M.A., Boutaba R., and Al-Kiswany S., "An approach for detecting multi-institution attacks," Annales des Telecommunications/Annals of Telecommunications, vol. 79, no. 3-4, pp. 257–270, 2024. doi: 10.1007/s12243-023-00993-4.

[24] Janabi A.H., Kanakis T., and Johnson M., "Convolutional Neural Network Based Algorithm for Early Warning Proactive System Security in Software Defined Networks," IEEE Access, vol. 10, pp. 14301–14310, 2022. doi: 10.1109/ACCESS.2022.3148134.

[25] Dimitriadis A., Lontzetidis E., Kulvatunyou B., Ivezic N., Gritzalis D., and Mavridis I., "Fronesis: Digital Forensics-Based Early Detection of Ongoing Cyber-Attacks," IEEE Access, vol. 11, pp. 728–743, 2023. doi: 10.1109/ACCESS.2022.3233404.

[26] Abu-Nassar A.M. and Morsi W.G., "Early Detection of Cyber-Physical Attacks on Electric Vehicles Fast Charging Stations Using Wavelets and Deep Learning," IEEE Transactions on Industrial Cyber-Physical Systems, vol. 2, pp. 220–231, 2024. doi: 10.1109/TICPS.2024.3413605.

[27] Kim T. and Pak W., "Early Detection of Network Intrusions Using a GAN-Based One-Class Classifier," IEEE Access, vol. 10, pp. 119357–119367, 2022. doi: 10.1109/ACCESS.2022.3221400.

[28] Kumari M., Gaikwad M., and Chavan S.A., "A secure IoT-edge architecture with data-driven AI techniques for early detection of cyber threats in healthcare," Discover Internet of Things, vol. 5, no. 1, Art. no. 54, 2025. doi: 10.1007/s43926-025-00147-z.

[29] Kok S.H., Abdullah A., and Jhanjhi N.Z., "Early detection of crypto-ransomware using pre-encryption detection algorithm," Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 5, pp. 1984–1999, 2022. doi: 10.1016/j.jksuci.2020.06.012.

[30] Wang M., Yang N., and Weng N., "K-GetNID: Knowledge-Guided Graphs for Early and Transferable Network Intrusion Detection," IEEE Transactions on Information Forensics and Security, vol. 19, pp. 7147–7160, 2024. doi: 10.1109/TIFS.2024.3431932.

[31] Di Monda D., Montieri A., Persico V., Voria P., De Ieso M., and Pescape A., "Few-Shot Class-Incremental Learning for Network Intrusion Detection Systems," IEEE Open Journal of the Communications Society, vol. 5, pp. 6736–6757, 2024. doi: 10.1109/OJCOMS.2024.3481895.

[32] Alsuwat E., Solaiman S., and Alsuwat H., "Concept Drift Analysis and Malware Attack Detection System Using Secure Adaptive Windowing," Computers, Materials and Continua, vol. 75, no. 2, pp. 3743–3759, 2023. doi: 10.32604/cmc.2023.035126.

[33] Kushal S., Shanmugam B., Sundaram J., and Thennadil S., "Self-healing hybrid intrusion detection system: an ensemble machine learning approach," Discover Artificial Intelligence, vol. 4, no. 1, Art. no. 28, 2024. doi: 10.1007/s44163-024-00120-9.

[34] Chen Y., Cui M., Wang D., Cao Y., Yang P., Jiang B., Lu Z., and Liu B., "A survey of large language models for cyber threat detection," Computers and Security, vol. 145, Art. no. 104016, 2024. doi: 10.1016/j.cose.2024.104016.