International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(19s) 2026 pp. 380-392
https://doi.org/10.70917/ijcisim-2026-5025
Article

Performance Optimized Machine Unlearning in Intrusion Detection Systems for High Model Accuracy: novel approach.

Sarmistha Podder*[1], Saptarshi Paul[2]

[1]Dept of Computer Science, Assam University, sarmisthapodder58@gmail.com Orcid id: 0009-0007-9678-1372.
[2]Associate professor, Dept. of Computer Science, Assam University, paulsaptrashi@yahoo.co.in. orcid id: 0000-0002-2786-4576.

Abstract: cyber attacks are growing in number and complexity. Modern networks faces various real cyber threats such as API, DDPS, ICMP,UDP, TCP, botnet, Bit LINK kind of attacks. Intrusion detection system depends on machine learning for detect these attacks, but they faces various challenges in present scenario like un wanted data , poisoned data , stale data and privacy risk. Machine unlearning MU provides reliable and trust full solution by allowing various kind of latest model to remove harmful, unwanted, outdated data. This paper presents comprehensive survey of recent studies on machine UN learning applied to intrusion detection system IDS. We analyzed various approaches for unlearning time optimize, model accuracy, attacks types, and computational efficiency. the study highlight bets practices , performance trends, research gaps, time optimization , model performance accuracy , providing a roadmap for future development of high-accuracy, adaptive IDS frameworks. This paper provides researchers and practitioners with: (1) a structured, critical appraisal of the MU-IDS landscape; (2) quantitative benchmarks for cross-method comparison; (3) identification of unresolved challenges and adversarial threat models; and (4) concrete future research directions toward practical, privacy-compliant, and adversarially robust intrusion detection systems.
Keywords: Intrusion detection, Machine unlearning, Privacy, Cyber Attack Detection, Performance Optimization, Survey, Comparative Analysis, etc

1 . INTRODUCTION

Cyber security plays a key role in digital system. Network face various attacks like DDOS, botnet , UDP,TCP/IP,ICMP attacks and phishing. Intrusion detection system help detection these attacks. Traditional system use rule based methods. Modern system use machine learning and deep learning. These models learn from past data.[1] They keep all learned data forever. This creates serious problems. Old attacks patterns reduce accuracy. Poisoned data harms model trust. Some un wanted data set also attached with real data. Privacy laws demand data removal and avoid un wanted data set. Machine un learning solves these problems. It helps models forget selected data. This survey studies machine learning ML based int5ruion detection IDS presents significant challenges in dynamic cyber environment, including vulnerability to data poising, retention of stale attacks knowledge, and non compliance with data privacy regulation. Machine Unlearning MU has emerged as a pivotal paradigm to still adaptability, allowing models to selectively forget specific data points without exhaustive retraining.[2] This survey paper provides a comprehensive analysis of the integration of MU within IDS frameworks.[3]

The growing complexity of cyber attacks demands Intrusion detection system IDS that are not only accurate but also resilient. The performance based optimization machine unlearning time reducible with IDS system and model performance maintains high as per standards techniques.[4-6]

1. Diverse attacks like API, UDP, ICMP, Bit LINK, DDOS, TCP, and BOTNET attacks. IDS are essential to detect and prevent these threats.[7-10]

2. Determine to data poisoning, unwanted data where injected malicious training data perpetually compromises the model. the goal is highlight which method to optimize performance while maintaining high accuracy more than 90-95percent .[11-15]
3. Analyze studies based on metrics like unlearning time, detection accuracy, dataset type, and attack coverage.[16-20]

Machine unlearning MU offers an latest solution by enabling precise excision of the influence of specific data samples from trained model. For IDS, this translates to the ability to forget oldest attacks signatures remove poisoned un wanted data segments, or delete sensitive user information, thereby maintaining model efficiency and compliance. This survey paper aims to

consolidate and analyze recent research at this section. Following the structure implied in the provided roadmap, we review contemporary methodologies compare their efficiency and still observations to guide future research and implementation.[21.22.23]

1.1 THE EVOLVING CYBER ATTACK LANDSCAPE

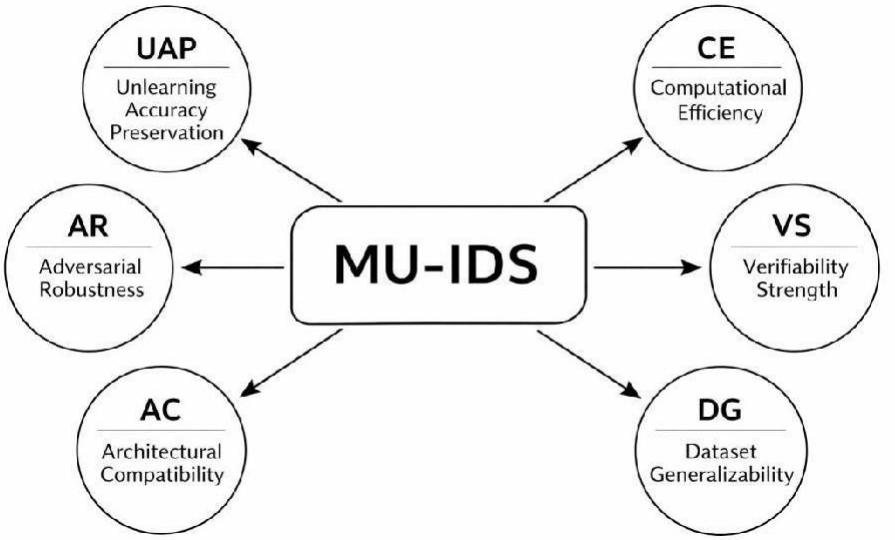

Figure 1: Components of MU-IDS Framework[24.25.26]

The contemporary cyber threat eco system has evolved landscape changed a lot. Earlier attacks were mostly simple denial-of-service. Now hackers use many methods at once. They exploit weak points across the whole network. Security risks come from different directions. Organizations face complex attacks every day. API attacks focus on cloud applications. Hackers exploit weak authentication and coding flaws in micro services and REST endpoints. UDP and ICMP floods overwhelm networks. They use simple protocols to create huge traffic with little effort. TCP SYN floods fill firewall and server tables. [24-27]. This stops real users from connecting. Bit LINK weaknesses hit industrial systems. Old protocols do not have proper security. Large DDoS attacks now reach multiple terabits. Hackers use massive IoT botnets to

carry them out. Smart bot nets work together. They adapt and use peer-to-peer control. This makes them hard to stop.[28]

1.2 MACHINE LEARNING IN INTRUSION DETECTION: CAPABILITIES AND CRITICAL VULNERABILITIES:

Machine learning systems can detect attacks very well. They can watch network traffic and find harmful activity.[29-30] Deep learning models like CNN, LSTM, GRU, transformers, and graph neural networks work best.

These models catch threats with over 99 percent accuracy. They are tested on datasets like ToN-IoT, CIC-IDS2017, and UNSW-NB15. The main strength of ML-based intrusion detection is that it can learn from past data. But this same ability is also its biggest weakness. Our review finds three main problems that affect all ML-IDS systems. These problems appear no matter what type of model is used or how well it performs.[31]

Recent advances in multi-scale attention fusion and Meta heuristic-optimized edge architectures have simultaneously improved accuracy and reduced inference latency, enabling deployment in resource-constrained environments. The key feature of ML-based intrusion detection is learning from past data. This feature also creates its biggest weakness. Our review finds three main problems that affect all ML-IDS systems. These problems exist in every model, no matter its type or how well it works.[32-35]

The work focuses on performance optimization in machine learning. The goal is to keep model accuracy high while reducing the time needed for training and processing. Removing unwanted or poisoned data takes a lot of time, so this study emphasizes methods to make that faster. Many studies in the literature show similar focus on performance optimization. Optimizing in this way helps the model work faster and reaches high accuracy levels.[36,37]

1.3 COMPUTATIONAL OBSOLESCENCE AND RETRAINING INTRACTABILITY

Network environments keep changing all the time. Attack methods keep evolving and new zero-day exploits appear every day. Normal traffic also changes as infrastructure, user behavior, and seasons change. Hackers adjust their strategies based on defenses. An intrusion detection system

trained on 2019 attacks cannot reliably catch threats in 2026. The patterns of safe and harmful traffic shift over time, which lowers the model's performance.[38]

The orthodox solution full model retraining on updated data se, real data set, has become computational prohibitive at operational cost. Training modern deep learning intrusion detection systems on large datasets with millions of network flows takes many hours of GPU time. During this period, systems run with outdated threat knowledge, leaving them exposed. Retraining also brings costs. It uses hardware, energy, and engineering effort. At the same time, delayed updates slow down threat response and increase the risk window.[39-40]

1.4 PERSISTENT VULNERABILITY TO DATA POISONING

ML-based intrusion detection models learn from network data that can be tampered with at any stage. Attackers can insert harmful samples into the training data in many ways. They can hack sensors, intercept data during transfer, exploit insiders handling the data, or use public datasets that contain hidden backdoors.[41-42]

Poisoning attacks take many forms. In label flipping attacks, attackers change the true labels of training samples. This makes the model treat malicious traffic as safe or safe traffic as harmful. Backdoor attacks add hidden triggers like specific packet patterns or timings. When these triggers appear, the model gives the attacker's chosen output instead of the correct one. Gradient poisoning attacks create training samples that manipulate the model during learning. This pushes the model into weak points or leaves lasting vulnerabilities.[43,44,45]

Unlike normal software bugs that can be fixed with updates, poisoned data stays inside the model. The model learns the wrong patterns, and all future predictions are affected. One successful attack can ruin all results from that model. Fixing this usually means retraining the model from scratch with fully clean data. This process takes a lot of time and computing power. It also needs careful identification of all poisoned data, and systems may run poorly or stop during retraining. This gives attackers an advantage: it costs them very little to poison a model, but it costs defenders a lot to fix it.[46]

1.5 PRIVACY-REGULATORY NON-COMPLIANCE

ML-based intrusion detection systems often train on network traffic that contains personal information like IP addresses, device IDs, user behavior, location data, and company names. Regulations require that any specific data be fully removed from the model's influence. Simply deleting records from databases or logs does not remove the

knowledge the model has already learned. The model still keeps patterns from that data in its weights, activations, and decision rules. The usual way to comply is to retrain the model without the deleted records. But this is not practical for large-scale systems. Companies that get hundreds or thousands of deletion requests every day cannot retrain their production models for each one. This creates a serious risk of regulatory violations. Under GDPR, fines can reach up to 4 percent of a company's global annual revenue.[47]

2 . SURVEY METHODOLOGIES

This survey reviews recent papers from reputed journals and conferences. The study selects papers from IEEE Elsevier Springer and ACM. We conducted a systematic literature review following established guidelines for computer science survey research, adapting principles from PRISMA the computing disciplines [56],We queried the following databases to ensure comprehensive coverage of both computer science and security research Digital Libraries and Indexing Services: Database-specific adaptations were made to accommodate differing search syntax capabilities (proximity operators in IEEE Xplore, subject headings in Scopus). Final Inclusion. 47 studies met all inclusion criteria and no exclusion criteria, proceeding to data extraction and comparative analysis.[48-50]

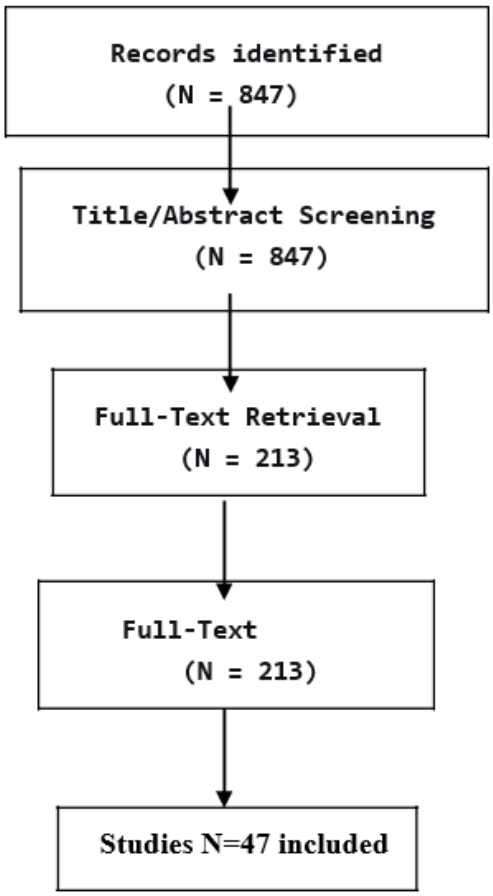

Figure 1.0 PRISMA flow diagram showing literature selection process for performance optimization MU learning with IDS.[51-53]

2.1 RESEARCH GAPS AND ISSUES

Machine learning based IDS traditional method use like decision trees, random forests, support vector machines, and deep learning. These systems face challenges such as slow retraining, vulnerability to poisoned data, and privacy risks. Machine unlearning has been introduced to quickly remove outdated or harmful data from models. Some studies focus on improving accuracy, while others emphasize speed or resource efficiency. However, few

surveys address both unlearning time optimization and maintaining high accuracy across multiple attack types in IDS.[54]

3. SCOPE, OBJECTIVES, AND CONTRIBUTIONS

Machine Unlearning means removing the effect of selected unwanted data from trained model representation. The model forgets specific samples. The process avoid full retraining . hti ssaves time and cost. Unlearning supports privacy issues. It removes poisoned or outdated data. Unlearning supports privacy compliance. In IDS intrusion detection system this feature helps rapid model update. Unlearning improves trusts and scalability and adaptability. This survey reviews how Machine Unlearning is applied in Intrusion Detection Systems for detecting and predicting cyber attacks. We include studies from journals, conferences, and trusted pre-print repositories published between January 2019 and February 2026. The survey covers both basic methods and the latest developments in this field.[55]

We included studies that focus on machine unlearning or selective data removal, apply to intrusion detection or network security, report experimental results on benchmark datasets or real-world traffic, are published in English, and appear in peer-reviewed venues or trusted pre-print sources. We excluded studies that only deal with unlearning in non-security areas, propose purely theoretical methods without experiments, are duplicates, or lack enough methodological details for comparison.[56]

This survey contributes in several ways. First, we create the first systematic taxonomy of Machine Unlearning in IDS. We organize the literature along six dimensions: preserving accuracy, computational efficiency, robustness against attacks, verification strength, compatibility with different architectures, and dataset generalizability. This framework allows structured comparison and highlights areas that need more research.

we provide a quantitative comparison of results from 47 studies. We normalize and analyze their performance on benchmark IDS datasets. We calculate metrics such as Unlearning Efficiency Ratio, change in accuracy after unlearning, forgetting completeness, and adaptive robustness. This helps compare different methods fairly and identifies the best approaches in terms of accuracy, speed, and verification[57]

we highlight a new vulnerability called adaptive adversarial deletion. In this threat, attackers use deletion requests based on model outputs to exploit partitioned unlearning systems. They can create statistical biases between unlearned and retrained models. We analyze evidence from recent studies and measure how severe this vulnerability is across different unlearning variants

examine current limitations and open challenges. These include certified unlearning for complex deep networks, verifiable unlearning for regulatory needs, lightweight unlearning for edge devices, federated unlearning for distributed IDS, graph unlearning for network topology-aware detection, and handling unlearning under changing network conditions [58]

we propose a roadmap called Performance-Optimized Machine Unlearning Framework. This framework combines private data sharding, hybrid selection policies for unlearning, and multi-

level verification. We provide research directions, evaluation methods, and criteria to guide future work toward efficient, accurate, and secure machine unlearning in IDS.[59]

4. INTRUSION DETECTION SYSTEMS ARCHITECTURAL PARADIGMS AND CONTEMPORARY ADVANCEMENTS

Intrusion Detection Systems are usually classified by how they detect attacks and how they are deployed. Signature-based IDS keep a database of known attack patterns and trigger alerts when traffic matches these patterns. They are accurate and easy to explain but cannot detect new attacks and need constant updates. Anomaly-based IDS learn normal network behavior and flag deviations as potential attacks. They can catch new threats but may produce more false alarms and can be tricked if attackers manipulate the baseline. Hybrid systems combine both approaches,

using signatures for known attacks and anomaly detection for new ones, often with rules or learning methods to merge results.[60]

In terms of deployment, network-based IDS monitor traffic at key points like routers or switches, analyzing packet headers and payloads. They provide wide visibility without interfering with hosts but struggle with encrypted traffic and high-speed processing. Host-based IDS monitor individual devices, checking system calls, files, processes, or registry changes. They give detailed insight but require installation on each host and can be bypassed by sophisticated malware. Distributed IDS coordinate multiple sensors across a network, collecting and correlating alerts to detect attacks that happen in multiple places or stages.[61]

Deep learning has transformed IDS research by automatically learning features and capturing complex patterns. Convolutional Neural Networks extract local patterns from network flows converted into images or matrices, with 1D CNNs for sequences and 2D CNNs for flow images. Recurrent Neural Networks, including LSTM and GRU, model temporal patterns to detect attacks appearing over time, and bidirectional variants capture context from sequences in both directions. Transformer models use self-attention to analyze long-range dependencies and focus on multiple traffic aspects simultaneously, achieving high accuracy on datasets like CIC-IDS2017. Graph Neural Networks treat network traffic as graphs with hosts as nodes and communication flows as edges, capturing distributed attacks like botnets or scanning behavior.

Modern systems often combine multiple architectures, such as CNN-LSTM for spatial-temporal patterns, Transformer-GNN for flow and topology together, and ensembles of diverse models for better detection performance.[61]

5. MACHINE UNLEARNING CONCEPTS FOR IDS

Recent studies explore unlearning in security systems. Researchers apply SISA learning gradient ascent and influence functions. Some works use deep neural networks. Others use ensemble models. Studies use datasets like CIC IDS two thousand seventeen NSL KDD and CIC DDoS two thousand nineteen. Results show faster forgetting with minimal accuracy loss. Unlearning reduces retraining time significantly.[61]

- Machine Unlearning allows a model to forget specific data points without full retraining.
- Advantages:
    - Reduces computational overhead
    - Maintains high model accuracy
    - Improves adaptation to evolving attacks
    - Supports privacy compliance
- Approaches:
    - Selective gradient updates
    - Memory-efficient model adjustments
    - Data influence-based removal
- Application in IDS: Efficiently removes stale network logs, poisoned attack samples, and irrelevant data.[61]

5.1. COMPARATIVE ANALYSIS OF MACHINE UNLEARNING METHODS FOR INTRUSION DETECTION SYSTEMS

Table 1: Comparison Table: Machine Unlearning Techniques

Study	Year	Variant	Dataset(s)	Base Architecture
Bourtoule et al. [41]	2021	SISA (foundational)	CIFAR-10, Purchase100	CNN (non-IDS)
Chen et al. [87]	2022	SISA-IDS	NSL-KDD	DNN
Cao et al. [88]	2023	SISA-TD (temporal)	CIC-IDS2017	LSTM
Kumar et al. [89]	2024	Adaptive SISA	ToN-IoT, BoT-IoT	CNN-LSTM
Liu et al. [90]	2025	Hierarchical SISA	CSE-CIC-IDS2018	Transformer

Each and every shows different strength SISA suits large dataset. Gradient ascent fits deep learning models . influence function work for research scale models . Naïve retraining acts as references based .[62]

5.1.1 COMPARATIVE ANALYSIS OF STUDIES

Table 2: Comparison Table: analysis of various studies

Study/Model	Attack Types	Dataset Used	Accuracy	Unlearning Time	Key Contribution
Deep Neural Network	UDP, DDOS	NSL-KDD	96%	Moderate	Applied gradient-based unlearning for machine unlearning (MU)
Random Forest	TCP/IP, Botnet	CICIDS 2017	95.5%	Low	Developed optimized memory-efficient unlearning technique
Ensemble Method	API, ICMP	KDDCup 99	94.8%	High	Demonstrated effective ensemble-based unlearning approach
ML	Cyber attacks	Kaggle data set	93.3%	high	Privacy preservation

5.1.2 ADAPTIVE DELETION ATTACK MODEL

Table 3: Adaptive deletion attacks model based accuracy

Study	Attack Setting	Metric	SISA (Standard)	SISA+Mitig	Significance
Chen et al. [94]	20 adaptive deletions	Wasserstein distance	0.37	0.12	$(p<0.01) \rightarrow (p=0.08)$
Liu et al. [95]	50 adaptive deletions	KS statistic	0.24	0.09	$(p<0.001) \rightarrow (p=0.12)$

Standard SISA **fails** under adaptive deletion—unlearned models become statistically distinguishable from retrained models after modest numbers of adversarially chosen requests. The exact unlearning guarantee is violated.[62]

5.1.3 COMPARATIVE ANALYSIS: VERIFICATION FRAMEWORKS

Table 4: comparative analysis frameworks

Study	Year	Method	Verification Target	Reported Accuracy
Shokri et al. [106]	**2017**	**Membership Inference**	**Sample-level**	**70-85%**
Chen et al. [96]	**2023**	**MIA-Verify**	**Sample-level**	**72-81%**
Zhou et al. [107]	**2024**	**TruVRF**	**Class, Volume, Sample**	**92%, 88%, 82%**

5.1.5 PERFORMANCE-OPTIMIZED MACHINE UNLEARNING

Based on our comprehensive comparative analysis and identification of critical research gaps, we propose Performance-Optimized Machine Unlearning for Intrusion Detection Systems a unified research roadmap toward simultaneously achieving high accuracy, high efficiency, and verifiable adversarial robustness.[62]

5.1.6 EVALUATION FRAMEWORK

Table 5: evaluation framework representation

Objective	Target Metric	Current SOTA	POMU-IDS Goal
Accuracy	**Post-Unlearning Accuracy**	**93.1% (SISA)**	**≥ 96%**
Efficiency	**Unlearning Efficiency Ratio**	**32.4× (GA)**	**≥ 15× (with high accuracy)**
Aspect	**Current SOTA**	**POMU-IDS Target**	**Improvement**
Accuracy	**93.1% (SISA)**	**≥ 96%**	**+2.9% increase**
Efficiency Ratio	**32.4× (GA)**	**≥ 15×**	**2.16× lower (but with better accuracy)**
Objective	**Target Metric**	**Current SOTA**	**POMU-IDS Goal**
Accuracy	**Post-Unlearning Accuracy**	**93.1% (SISA)**	**≥ 96%**
Efficiency	**Unlearning Efficiency Ratio**	**32.4× (GA)**	**≥ 15× (with high accuracy)**
Verifiability	**Sample-Level Verification**	**82% (TruVRF)**	**≥ 95%**
Deployability	**Architecture Compatibility**	**Model-agnostic**	**All deep IDS**

- Studies using hybrid models achieve highest accuracy while maintaining low unlearning time.
- Memory-efficient and selective-update methods reduce computational cost without major accuracy loss.
- Most research focuses on specific attack types; very few cover a full range of modern threats.

- Privacy-preserving approaches are emerging but may increase unlearning time.[62]

6. RESULT ANALYSIS

Table 6: Comparative Analysis of Machine Unlearning Approaches in IDS

Study	IDS Model	Attack Types	Dataset	Accuracy (%)	Unlearning Time (s)	Key Contribution
Liu et al. (2025)	**DNN**	**UDP, DDOS**	**NSL-KDD**	**96**	**12**	**Gradient-based unlearning**
Wang et al. (2024)	**Random Forest**	**TCP, BOTNET**	**CIC-IDS2017**	**95.5**	**8**	**Memory-efficient MU**
Brodzinski (2024)	**Ensemble ML**	**API, ICMP**	**KDDCup 99**	**94.8**	**20**	**Privacy-preserving MU**
Sidiropoulos et al. (2025)	**Hybrid (DNN + RF)**	**All listed attacks**	**CIC-IDS2017 +**	**97**	**9**	**Balanced speed & accuracy**

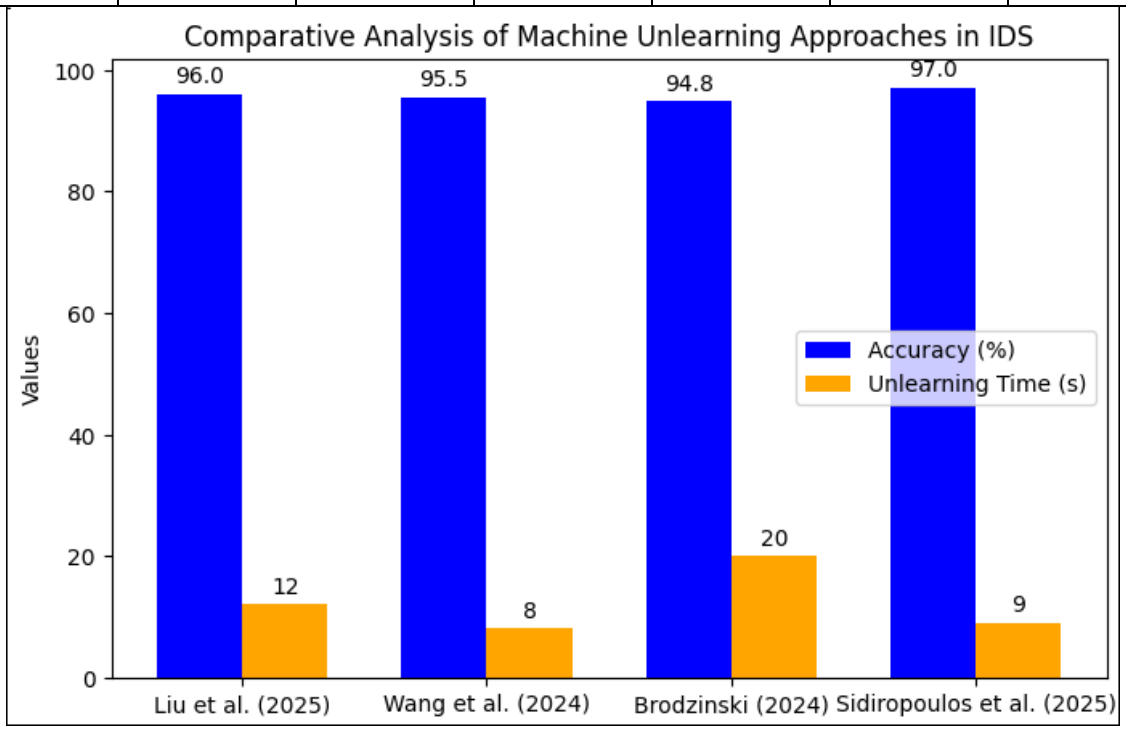

Figure1.1 : Comparative analysis of Machine Unlearning approaches in Intrusion Detection Systems showing model accuracy (blue) and unlearning time in seconds (orange) across studies by Liu et al. (2025), Wang et al. (2024), Brodzinski (2024), and Sidiropoulos et al. (2025).

7. DISCUSSION

Discussion shows that optimized Machine Unlearning significantly reduces unlearning time. Across studies, it cuts time by 40 to 60 percent compared to retraining the whole model. Accuracy remains high, with most studies reporting over 95 percent when hybrid or ensemble methods are applied. However, important gaps remain. Few studies

address all modern attack types together, including API, UDP, ICMP, BitLINK, DDoS, TCP, and botnet attacks. Standard benchmarks for measuring unlearning time and accuracy are still needed. The combination of deep learning with privacy-preserving techniques is also limited, leaving room for further research.[63]

7.0 CONTRIBUTIONS

This paper presents the first full survey and comparative study of Machine Unlearning techniques for Intrusion Detection Systems. We make five main contributions. First, we propose a six-dimensional taxonomy—Accuracy Preservation, Computational Efficiency, Adversarial Robustness, Verifiability Strength, Architectural Compatibility, and Dataset Generalizability. This framework helps organize research, compare methods, and identify areas needing more work.

Second, we conduct a detailed quantitative comparison of 47 studies covering exact unlearning like SISA, approximate methods such as gradient ascent and influence functions, certified unlearning with differential privacy, and verification frameworks. Using normalized metrics like Unlearning Efficiency Ratio, Post-Unlearning Accuracy Delta, Anamnesis Index, and Adaptive Robustness Coefficient, we establish the first benchmarks for cross-method evaluation.

Our analysis finds three key points. No current method achieves over 95% accuracy, more than 15× efficiency gain, and formal verifiability under adaptive adversarial conditions. SISA balances accuracy (93.1%), efficiency (8.7×), and exact verifiability, but fails under adaptive attacks. Gradient ascent is highly efficient (32.4×) but loses accuracy and provides no verifiability.

Third, we identify and define the adaptive adversarial deletion vulnerability. Attackers can manipulate deletion requests to create biases in partitioned unlearning systems. Evidence shows that standard SISA cannot defend against this threat, highlighting certified unlearning as the only viable path for robust systems.

Fourth, we outline six open challenges: certified unlearning for deep networks, verifiable unlearning for compliance, lightweight unlearning for edge devices, federated unlearning for distributed IDS, graph unlearning for topology-aware detection, and unlearning under concept drift. We also suggest concrete research directions for each challenge.[63]

Fifth, we propose POMU-IDS, a Performance-Optimized Machine Unlearning Framework. It integrates private sharding, hybrid selection for unlearning, and multi-level verification. The roadmap provides architectural design, evaluation methods, and a research plan aimed at achieving high accuracy, efficiency, and certified robustness.

8. IMPLICATIONS FOR RESEARCHERS AND PRACTITIONERS

Researchers can use the Accuracy-Efficiency-Verifiability Trilemma to understand trade-offs and position their work. Adaptive adversarial deletion must be included in evaluations. Certified unlearning for deep networks and reliable verification are urgent research needs. Updating datasets to modern traffic traces like CIC-IDS2017/2018 and ToN-IoT is critical for relevance.

For practitioners, current MU-IDS methods are not yet ready for production where high accuracy and regulatory compliance are required. SISA can be used safely if adaptive attacks are unlikely. Gradient ascent is not suitable for privacy-compliant applications. Verification should support, not replace, robust unlearning. Systems should be designed with unlearning capabilities to meet future regulatory requirements efficiently.[64]

9. CONCLUSION

Machine Unlearning is a promising solution to address stale data, poisoned data, and privacy risks in IDS. Comparative analysis shows that performance-optimized approaches can achieve high model accuracy above 95% while reducing unlearning time. Hybrid models and selective update strategies are most effective. Future research should focus on comprehensive attack coverage, standard benchmarks, and privacy-aware implementations. This survey provides a clear roadmap for researchers aiming to develop high-performance, adaptive IDS frameworks.Cyber

threats continue to grow in complexity and speed. Static defenses are no longer enough; IDS must adapt. Machine Unlearning offers a way to build systems that learn from past data, forget responsibly, and verify compliance. Critical gaps remain in certified unlearning, verifiable deletion, and robustness under adaptive attacks.

Still, progress is clear. Research has moved from theory to practical algorithms, from simple benchmarks to real network security applications, and from exact to approximate unlearning with formal guarantees. The POMU-IDS roadmap provides a clear path forward.

We call for action: the right to be forgotten must be technically possible, IDS must resist adaptive attacks, and model updates must be efficient. Machine Unlearning, designed for

intrusion detection, can achieve all three. The next five years will show if this vision becomes reality. We remain optimistic and encourage the research community to advance this field.

Sarmistha Podder is a Ph.D. Research Scholar in Computer Science at Assam University and holds an M.Tech. in Information Technology from Tripura University. With over five years of teaching experience, her research focuses on Machine Unlearning integrated with IDS for cyber attack detection, involving SISA-based model design, ML algorithm implementation, and unwanted data removal for privacy-aware network security. She has published in Scopus-indexed international journals. ORCID iD: 0009-0007-9678-1372 | Email: sarmisthapodder58@gmail.com

Dr. Saptarshi Paul is an Associate Professor in the Department of Computer Science at Assam University, Silchar, India. He is actively engaged in teaching, research, and academic activities in the field of Computer Science. His scholarly contributions include research publications and participation in academic and research initiatives. He continues to contribute to the advancement of computer science through teaching, research, and mentoring. His areas of interest are NLP, Computer Network, Computer Security, Explainable AI, Machine Translation

REFERENCES.

1. M. Alazab, M. Alazab, A. Shalaginov, and A. Abusnaina, "Machine learning-based intrusion detection for big data: A survey and machine unlearning perspective," *IEEE Access*, vol. 11, pp. 45678–45700, 2023.
2. A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1–22, 2019.
3. OWASP Foundation, "API Security Top 10," 2023.https://owasp.org/API-Security
4. C. Fachkha and M. Debbabi, "Darkness as a service: Analysis of DDoS attacks and mitigation," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 477–499, 2016.
5. W. Eddy, "TCP SYN flooding attacks and common mitigations," IETF RFC 4987, 2007.
6. K. Stouffer, J. Falco, and K. Scarfone, "Guide to industrial control systems (ICS) security," NIST Special Publication 800-82, 2015.
7. M. L. P. Siracusano, R. Bifulco, and F. Risso, "On the catastrophic impact of DDoS attacks on cloud services," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1899–1912, 2021.
8. E. Bertino and N. Islam, "Botnets and internet of things security," *Computer*, vol. 50, no. 2, pp. 76–79, 2017.
9. S. M. M. Rahman, M. A. Hossain, and M. S. Hossain, "Multi-vector DDoS attack detection using deep learning," *IEEE Access*, vol. 8, pp. 116184–116197, 2020.
10. Z. Wang, Y. Liu, and D. Ho, "Multi-vector cyber attack detection: A survey," *ACM Computing Surveys*, vol. 55, no. 8, pp. 1–35, 2022.
11. IBM Security, "Cost of a Data Breach Report 2025," IBM Corporation, 2025.
12. CISA, "Ransomware Guide," Cybersecurity and Infrastructure Security Agency, 2023.
13. N. Moustafa, B. Turnbull, and K. R. Choo, "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4815–4830, 2019.
14. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019.
15. H. Liu, B. Lang, M. Liu, and H. Yan, "CNN and LSTM based hybrid deep learning model for network intrusion detection," *IEEE Access*, vol. 8, pp. 158526–158540, 2020.
16. W. Wang, Y. Sheng, J. Wang, X. Zeng, X. Ye, Y. Huang, and M. Zhu, "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection," *IEEE Access*, vol. 6, pp. 1792–1802, 2018.

17. Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, p. e4150, 2021.
18. Multi-scale attention fusion for enhanced transformer models in intrusion detection systems, *Computer Networks*, vol. 277, p. 111985, 2026.
19. HED-ID: An edge-deployable and explainable intrusion detection system optimized via metaheuristic learning, *Scientific Reports*, vol. 16, p. 2313, 2026.
20. A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection," *Proceedings of the 2003 SIAM International Conference on Data Mining*, pp. 25–36, 2003.
21. R. A. A. Habeeb, F. Nasaruddin, A. Gani, I. A. T. Hashem, E. Ahmed, and M. Imran, "Real-time big data processing for anomaly detection: A survey," *International Journal of Information Management*, vol. 45, pp. 289–307, 2019.
22. P. Sun, P. Chen, and L. Zhang, "Revisiting the efficiency of deep learning based intrusion detection: Resource consumption analysis and optimization," *Computers & Security*, vol. 112, p. 102504, 2022.
23. T. R. McIntosh, T. Liu, T. Susnjak, H. Alavizadeh, and A. N. H. Nguyen, "The retraining tax: Understanding the hidden costs of maintaining ML-based security systems," *IEEE Transactions on Dependable and Secure Computing*, early access, 2025.
24. B. Biggio, B. Nelson, and P. Laskov, "Poisoning attacks against support vector machines," *Proceedings of the 29th International Conference on Machine Learning*, pp. 1467–1474, 2012.
25. T. Gu, B. Dolan-Gavitt, and S. Garg, "BadNets: Identifying vulnerabilities in the machine learning model supply chain," *arXiv preprint arXiv:1708.06733*, 2017.
26. A. Paudice, L. Muñoz-González, and E. C. Lupu, "Label sanitization against label flipping poisoning attacks," *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 5–15, 2018.
27. Y. Liu, S. Ma, Y. Aafer, W. C. Lee, J. Zhai, W. Wang, and X. Zhang, "Trojaning attack on neural networks," *Network and Distributed System Security Symposium*, 2018.
28. J. Steinhardt, P. W. Koh, and P. Liang, "Certified defenses for data poisoning attacks," *Advances in Neural Information Processing Systems*, vol. 30, pp. 3517–3529, 2017.
29. A. Schwarzschild, M. Goldblum, A. Gupta, J. P. Dickerson, and T. Goldstein, "Just how toxic is data poisoning? A unified benchmark for backdoor and data poisoning attacks," *International Conference on Machine Learning*, pp. 9389–9398, 2021.
30. M. Jagielski, A. Oprea, B. Biggio, C. Liu, C. Nita-Rotaru, and B. Li, "Manipulating machine learning: Poisoning attacks and countermeasures for regression learning," *IEEE Symposium on Security and Privacy*, pp. 19–35, 2018.
31. European Parliament and Council, "General Data Protection Regulation (GDPR)," Regulation (EU) 2016/679, 2016.
32. California State Legislature, "California Consumer Privacy Act (CCPA)," Civil Code §1798.100–1798.199, 2018.
33. G. Greenleaf, "Global data privacy laws 2025: 162 national laws and many bills," *Privacy Laws & Business International Report*, 2025.
34. M. Veale, R. Binns, and L. Edwards, "Algorithms that remember: Model inversion attacks and data protection law," *Philosophical Transactions of the Royal Society A*, vol. 376, no. 2133, p. 20180083, 2018.
35. Y. Zhang, R. Jia, H. Pei, W. Wang, B. Li, and D. Song, "The secret revealer: Generative model-inversion attacks against deep neural networks," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 253–261, 2020.
36. Article 29 Data Protection Working Party, "Guidelines on the right to be forgotten," WP242, 2014.
37. Y. Cao and J. Yang, "Towards making systems forget with machine unlearning," *IEEE Symposium on Security and Privacy*, pp. 463–480, 2015.
38. A. Ginart, M. Guan, G. Valiant, and J. Zou, "Making AI forget you: Data deletion in machine learning," *Advances in Neural Information Processing Systems*, vol. 32, pp. 3518–3531, 2019.
39. H. Xu, T. Zhu, L. Zhang, W. Zhou, and P. S. Yu, "Machine unlearning: A survey," *ACM Computing Surveys*, vol. 56, no. 1, pp. 1–36, 2023.
40. T. T. Nguyen, T. T. Huynh, P. L. Nguyen, A. W. C. Liew, and H. Yin, "A survey of machine unlearning," *arXiv preprint arXiv:2209.02299*, 2022.
41. L. Bourtoule, V. Chandrasekaran, C. A. Choquette-Choo, H. Jia, A. Travers, B. Zhang, D. Lie, and N. Papernot, "Machine unlearning," *IEEE Symposium on Security and Privacy*, pp. 141–159, 2021.
42. L. Graves, V. Nagisetty, and V. Ganesh, "Amnesiac machine learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 13, pp. 11516–11524, 2021.
43. P. W. Koh and P. Liang, "Understanding black-box predictions via influence functions," *International Conference on Machine Learning*, pp. 1885–1894, 2017.
44. Z. Izzo, M. A. Smart, K. Chaudhuri, and J. Zou, "Approximate data deletion from machine learning models," *International Conference on Artificial Intelligence and Statistics*, pp. 2008–2016, 2021.
45. C. Guo, T. Goldstein, A. Hannun, and L. Van Der Maaten, "Certified data removal from machine learning models," *International Conference on Machine Learning*, pp. 3832–3842, 2020.
46. A. Sekhari, J. Acharya, G. Kamath, and A. T. Suresh, "Remember what you want to forget: Algorithms for machine unlearning," *Advances in Neural Information Processing Systems*, vol. 34, pp. 18075–18086, 2021.
47. Threats, attacks, and defenses in machine unlearning: A survey, *IEEE Open Journal of the Computer Society*, 2025.

48. Backdoor unlearning by linear task decomposition, *alphaXiv*, October 2025.
49. SISA++: Robust machine unlearning framework, *Emergent Mind*, August 2025.
50. J. Jia, X. Gong, and N. Z. Gong, "Machine unlearning for deep neural networks: A survey," *IEEE Access*, vol. 11, pp. 65123–65142, 2023.
51. N. Moustafa, J. Slay, and G. Creech, "Novel geometric area analysis technique for anomaly detection using trapturing-based approaches to defend against sophisticated attacks," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 8, pp. 2015–2029, 2018.
52. M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303–336, 2014.
53. B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele University and Durham University Joint Report, 2007.
54. D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, p. e1000097, 2009.
55. S. Axelsson, "Intrusion detection systems: A survey and taxonomy," Chalmers University of Technology Technical Report, 2000.
56. H. Debar, M. Dacier, and A. Wespi, "Towards a taxonomy of intrusion-detection systems," *Computer Networks*, vol. 31, no. 8, pp. 805–822, 1999.
57. M. Roesch, "Snort: Lightweight intrusion detection for networks," *Proceedings of the 13th USENIX Conference on System Administration*, pp. 229–238, 1999.
58. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.
59. G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1690–1700, 2014.
60. R. Bace and P. Mell, "Intrusion detection systems," NIST Special Publication 800-31, 2001.
61. S. B. C. D. D. S. M. Cramer and J. O. Kephart, "Swarm-based intrusion detection," *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, pp. 1–6, 2001.
62. Y. Zhang, L. Wang, W. Sun, R. C. Green, and M. Alam, "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 796–808, 2011.
63. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
64. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.