TSARKA SCIENCE

Volume 1 (2026), Issue 2
pp. 2–19

eISSN: 3134-6057
pISSN: 3134-6049

DOI: 10.66571/tsarka-3134-6057-04

A SYSTEMATIC LITERATURE REVIEW ON MACHINE LEARNING MODELS FOR ACTIVE DIRECTORY ATTACK DETECTION USING WINDOWS SECURITY EVENT LOGS

A.A. Tastemirov[1]*, M.B. Bekbolatov[1]

[1]TSARKA, Astana, Kazakhstan
*Corresponding author: azamat.t@tsarka.com

Abstract

Active Directory (AD) serves as the foundational identity and access management infrastructure in the vast majority of enterprise Windows environments and has become the primary target of sophisticated cyberattacks, including Kerberoasting, Pass-the-Hash, Golden Ticket, DCSync, lateral movement, and Advanced Persistent Threat (APT) campaigns. Traditional signature-based detection mechanisms have demonstrated systematic inadequacy against stealthy, behaviorally adaptive adversaries who exploit legitimate protocol functionality to evade detection. This literature review systematically examines 30 peer-reviewed and technically validated works published between 2018 and 2025, specifically focused on the detection of threats in Active Directory environments using machine learning, deep learning, and hybrid model architectures. The reviewed works are analyzed across five dimensions: AD-specific attack taxonomies, Windows Security Event Log feature engineering, classical and deep learning detection algorithms, hybrid and ensemble architectures, and graph-based approaches integrating provenance analysis and attack graphs. Key findings indicate that no single algorithm achieves comprehensive coverage across the diversity of AD attack types; instead, hybrid architectures combining sequential modeling (LSTM, BiLSTM), graph-based analysis (GNN), and ensemble classification (Random Forest, XGBoost) demonstrate superior detection accuracy and practical deployability. Provenance-graph-based systems exhibit particular promise for detecting multi-stage APT campaigns in real AD environments. Five critical research gaps are identified: the absence of large-scale labeled AD-specific Security Event Log datasets, the lack of unified hybrid frameworks validated on AD authentication data, insufficient multi-stage kill-chain detection capability, unresolved real-time deployment challenges, and limited application of explainable AI techniques to AD security contexts. These gaps define the research agenda for the development of

Volume 1 (2026), Issue 2
pp. 2–19

TSARKA SCIENCE

eISSN: 3134-6057
pISSN: 3134-6049

novel hybrid machine learning models for AD-based threat prediction and detection.

Keywords: *Active Directory, threat detection, hybrid machine learning, deep learning, intrusion detection, Windows Event Log, Kerberoasting, lateral movement, provenance graph, ensemble learning, anomaly detection, SIEM, APT detection.*

1. Introduction

Microsoft Active Directory Domain Services (AD DS) underpins authentication, authorization, and directory management in the overwhelming majority of enterprise Windows environments. As the central identity plane of corporate infrastructure, AD has become the most strategically valuable and intensively targeted component in modern cyberattacks. A successful compromise of an AD domain controller grants an attacker unrestricted access to all domain resources - users, machines, services, and sensitive data - making AD security a problem of fundamental organizational importance.

Recent research has documented a marked increase in the sophistication of AD-targeting techniques. Opanovych [1] demonstrates that modern APT actors systematically exploit AD weaknesses through multi-stage kill chains involving reconnaissance, credential theft, lateral movement, and domain dominance - all while mimicking legitimate user and administrative activity to evade rule-based detection systems. Similarly, Matsuda et al. [2] establish through controlled experimentation that machine learning approaches significantly outperform traditional signature-based detection for AD-specific attack scenarios, motivating a broader research effort into ML-driven AD security.

The limitations of traditional defenses are particularly pronounced for attacks that operate within the bounds of legitimate protocols. Kerberoasting, documented in depth by Kotlaba et al. [3], exploits the standard Kerberos service ticket request mechanism to extract crackable credential hashes without generating readily detectable anomalies. Pass-the-Hash and Golden Ticket attacks similarly abuse legitimate NTLM and Kerberos authentication workflows, leaving only subtle statistical and behavioral signals that require ML-based anomaly detection to surface reliably.

The research community has responded with a diverse portfolio of detection approaches spanning classical ML, deep learning, graph-based analysis, and hybrid ensemble methods. Mabika [4] demonstrates the viability of supervised learning applied directly to AD logs for overcoming cybersecurity challenges. The HADES system [5] achieves reliable provenance-graph-based detection of AD attacks including Pass-the-Hash and Golden Ticket in real domain controller environments. Herranz-Oliveros et al. [6] employ unsupervised learning on AD attack graphs specifically for lateral movement threat mitigation. However, as Opanovych [1] notes, each individual approach exhibits distinct limitations: classification algorithms struggle with high

TSARKA SCIENCE

Volume 1 (2026), Issue 2
pp. 2–19

eISSN: 3134-6057
pISSN: 3134-6049

behavioral variability, while graph-based methods face scalability challenges at enterprise scale.

The objective of this literature review is to systematically synthesize the current research landscape at the intersection of Active Directory security and machine learning, with the specific aim of identifying the algorithmic foundations and architectural patterns most suitable for the development of a novel hybrid model for AD-based threat prediction and detection. The review covers 30 works published between 2018 and 2025, selected for their direct relevance to AD-specific detection challenges, and is organized as follows: Section 2 describes the research methodology; Section 3 presents results across seven thematic areas; the Conclusion identifies key research gaps and future directions.

2. Materials and Methods

The concepts of «competitiveness», «competition» change, deepen, take on a new form and new facets along with the development of world economic thought, reflecting its genesis over time. Thus, different authors identify different factors of competitiveness: resources and market environment [2], [3], innovation – [4, 5], choosing a successful strategy of coexistence [6] or creating their own spaces for competition, the so-called «blue oceans» – spaces where the company will be the only player or leader [7], ensuring proximity to consumers, taking into account their requirements, leadership in the development of new products [8, p. 85-91], formation of a balanced corporate culture and interaction with consumers [9, p. 42], knowledge management [10], business model [11], organizational culture [12] and others.

2.1 Search Strategy and Source Selection

Literature was retrieved from IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, ResearchGate, arXiv (cs.CR, cs.LG), and institutional repositories. The primary search query combined: (“Active Directory” OR “Kerberos” OR “domain controller”) AND (“machine learning” OR “deep learning” OR “anomaly detection” OR “intrusion detection” OR “hybrid model”) AND (“threat detection” OR “attack detection” OR “event log” OR “provenance graph”). Additional targeted queries were applied for specific attack types (Kerberoasting, lateral movement, APT, ransomware), specific model architectures (LSTM, GNN, ensemble, hybrid), and specific application domains (SIEM, UEBA, XDR). A total of 30 works were selected for inclusion based on their direct relevance to the research topic.

2.2 Inclusion and Exclusion Criteria

Works were included if they addressed at least one of: (a) ML or DL-based detection of AD-specific attacks; (b) supervised, unsupervised, or hybrid learning applied to Windows Security Event Logs or AD telemetry; (c) graph-based analysis of AD attack graphs or provenance data; (d) hybrid or ensemble model architectures applied

to cybersecurity detection tasks with demonstrable relevance to AD environments; or (e) frameworks for AI-assisted security assessment of AD environments. Works focused exclusively on network traffic without any AD-specific component were excluded, as were works without verifiable publication venues.

2.3 Analysis Framework

Each selected work was analyzed along five dimensions: (1) the specific AD attack types or anomaly categories addressed; (2) the data sources, feature engineering approaches, and datasets employed; (3) the ML or DL algorithms and architectural design; (4) quantitative performance results and evaluation methodology; and (5) practical deployment context including scalability, real-time capability, and SIEM integration. Findings were synthesized into seven thematic subsections in Section 3.

3. Results and Discussion

3.1 Taxonomy of Active Directory Attacks and Detection Signals

Analysis of the reviewed literature reveals a consistent taxonomy of AD-specific attacks organized by the protocol layer exploited and the adversarial kill chain phase. Table 1 presents the consolidated taxonomy derived from the most comprehensive threat analyses in the corpus.

Kerberoasting is among the most extensively studied AD-specific attacks in the ML detection literature. Kotlaba et al. [3] provide the most rigorous ML-based treatment, applying One-Class SVM and Local Outlier Factor (LOF) to Windows Event ID 4769 data from a real AD environment with hundreds of daily users. Their key finding is that ML-based anomaly detection significantly reduces the false-positive rate compared to static threshold rules - from 9.7% (threshold rule) to 2.3% (One-Class SVM) - while maintaining full true-positive detection. This result establishes the baseline for anomaly-based Kerberoasting detection. A complementary analysis is provided by the METU thesis [7], which extends Kerberoasting detection to supervised ML algorithms, examining a broader set of classifiers beyond the anomaly detection paradigm.

APT detection in AD environments is addressed by Matsuda et al. [2] and Opanovych [1]. Matsuda et al. [2] propose a multi-layer ML detection system for APT attacks against AD, demonstrating that behavioral features derived from authentication event sequences substantially improve detection accuracy over network-only approaches. Opanovych [1] conducts a systematic comparative evaluation of cluster, classifier, neural network, and graph-based algorithms on a simulated AD environment. The study finds LSTM models to be most effective at detecting anomalies in user behavior sequences, while graph-based algorithms are optimal for network traffic anomaly detection - a result that directly motivates the hybrid approach of the proposed research.

TSARKA SCIENCE

Volume 1 (2026), Issue 2
pp. 2–19

eISSN: 3134-6057
pISSN: 3134-6049

Lateral movement detection is studied by Uppstromer and Raberg [8] using supervised ML on AD log files. Their study compares multiple classifiers on a semi-synthetic AD dataset incorporating PtH, PtT, and AD enumeration activities, finding high accuracy across classifiers but significant variance in precision-recall trade-offs depending on the attack subtype. The HADES system [5] extends lateral movement detection to the provenance graph domain, detecting both PtH and Golden Ticket attacks through whole-system provenance analysis of AD events.

Table 1. Taxonomy of Active Directory attacks addressed in reviewed literature with detection approaches.

Attack Type	Key Event IDs	ML Detection Approach	Key Reference
1	2	3	4
Kerberoasting	4769 (RC4 encryption)	Anomaly detection on ticket request stats	[3] Kotlaba et al. 2021
Pass-the-Hash	4624 (Type 3), 4776	Provenance graph / sequential anomaly	[5] HADES 2025
Golden Ticket	4769, 4672, 4768	Provenance graph, ticket lifetime anomaly	[5] HADES 2025
APT Kill Chain	Multiple correlated	LSTM, Graph-based, multi-stage modeling	[1] Opanovych 2025
Lateral Movement	4624, 4648, 4688	Supervised ML on AD log files	[8] Uppstromer 2019
Data Exfiltration	Multiple (tactic seq.)	MITRE-tactic ML sequence prediction	[9] ARKAIV 2025
Ransomware in AD	Process + auth events	Behavioral detection, automated signature	[10] AD Ransomware 2024

3.2 Security Event Log Analysis and Feature Engineering

Effective feature engineering from Windows Security Event Logs is the critical prerequisite for ML-based AD threat detection. The reviewed literature identifies three primary feature engineering paradigms, each suited to different attack types and algorithmic approaches.

Statistical and behavioral aggregation features derive per-user, per-session, and per-time-window statistics from raw event streams. Kotlaba et al. [3] demonstrate this paradigm in the context of Kerberoasting, constructing features representing the number of distinct service ticket requests per source account per day, stratified by service type and account type. This aggregation approach effectively captures the bulk statistical signature of Kerberoasting while filtering the high-frequency noise of

normal ticket requests. Haq et al. [11] extend behavioral feature engineering to insider threat detection in AD logs, applying NLP word embedding techniques (Word2Vec, GloVe) to event message text to capture semantic content beyond numeric event IDs - a critical innovation because the textual content of security event messages carries discriminative information about the context and intent of the recorded action.

Tactic-level sequence features represent a higher-level abstraction, mapping low-level event logs to MITRE ATT&CK tactics and modeling the sequence of tactics as the detection input. Hakim et al. [9] introduce this paradigm in the ARKAIV system, which bridges the gap between low-level AD event logs and high-level conceptual frameworks for data exfiltration prediction. By transforming authentication and access events into tactic sequences and applying supervised ML to predict exfiltration occurrences, ARKAIV achieves detection that is invariant to specific tool signatures while remaining sensitive to the strategic intent of the attacker.

Graph-based features exploit the natural graph structure of AD - users, machines, groups, and services as nodes; authentication, replication, and access events as directed edges. The HADES system [5] constructs whole-system provenance graphs from AD event telemetry, applying graph analysis to detect attack patterns that manifest as anomalous sub-graphs. Nebbione and Calzarossa [12] represent AD infrastructure and its vulnerabilities as attack graphs, extracting attack paths as features for ML classification - an approach that achieves F1-score of 0.91 for Random Forest classification of vulnerable AD configurations.

3.3 Classical Machine Learning Approaches

Supervised classification algorithms applied to AD event features form the most extensively validated detection paradigm in the reviewed corpus. Mabika [4] applies multiple supervised classifiers to AD event log data, demonstrating that tree-based ensemble methods consistently achieve the highest accuracy-precision-recall trade-off for AD-specific attack scenarios. The supervised learning approach is particularly effective when applied to well-characterized attack types with abundant labeled examples, such as the authentication-layer attacks documented in the experimental AD environments of Uppstromer and Raberg [8].

Kotlaba et al. [3] provide the most controlled comparison of classical ML approaches for AD-specific detection. Their study evaluates One-Class SVM (novelty detection) and Local Outlier Factor (outlier detection) on real Kerberoasting data, finding that One-Class SVM achieves superior true-positive detection with a competitive false-positive rate of 2.3% versus 9.7% for threshold-based rules at equivalent sensitivity. This result establishes One-Class SVM as the strongest single-algorithm baseline for unsupervised Kerberoasting detection.

Nebbione and Calzarossa [12] apply multiple classical ML classifiers - including Random Forest, Logistic Regression, and Decision Tree - to the problem of AI-assisted

TSARKA SCIENCE
Volume 1 (2026), Issue 2
pp. 2–19
eISSN: 3134-6057
pISSN: 3134-6049

security assessment of AD environments, evaluated on 220 artificially generated AD configurations. Random Forest achieves F1-score of 0.91 for identifying vulnerable configurations, confirming its robustness for graph-structured AD feature classification. Haq et al. [11] compare XGBoost, AdaBoost, Random Forest, KNN, and Logistic Regression against deep learning LSTM models for insider threat detection in AD logs, finding that ML-based models outperform DL-based ones on their specific dataset - an important counterpoint to the generally held assumption that DL always surpasses classical ML.

Table 2. Comparative overview of classical ML approaches applied to AD threat detection.

Algorithm	AD Application	Key Result	Reference
1	2	3	4
One-Class SVM	Kerberoasting anomaly detection	2.3% FPR vs 9.7% rules	[3] Kotlaba 2021
Random Forest	AD vulnerability assessment	F1 = 0.91	[12] Nebbione 2023
XGBoost / RF	Insider threat in AD logs	Outperform LSTM/GloVe	[11] Haq 2022
Multiple classifiers	Lateral movement in AD logs	High acc., variable PR	[8] Uppstromer 2019
LOF	Kerberoasting outlier detection	Lower recall vs SVM	[3] Kotlaba 2021
Supervised ML	AD general supervised detection	Competitive baseline	[4] Mabika 2024

3.4 Deep Learning Approaches for AD Threat Detection

Deep learning approaches have been applied to AD security from three directions: recurrent sequence modeling of event logs, hybrid deep learning architectures for general security attack detection, and neural network-based dynamic defense of AD attack graphs.

Opanovych [1] provides the most comprehensive comparative evaluation of DL approaches specifically in AD environments, comparing LSTM, Autoencoder, and graph-based neural networks across multiple attack categories. The study finds that LSTMs are superior for detecting anomalies in user behavior sequences - including authentication and access patterns characteristic of APT lateral movement - while Autoencoders excel at detecting anomalies in command-line execution sequences.

Volume 1 (2026), Issue 2
pp. 2–19

TSARKA SCIENCE

eISSN: 3134-6057
pISSN: 3134-6049

Crucially, the study demonstrates that no single DL architecture achieves comprehensive detection across all AD attack types, establishing the empirical foundation for hybrid model development.

Hybrid deep learning models developed for general security contexts provide important architectural insights applicable to AD detection. Sagu et al. [13] propose a hybrid DL model with self-improved optimization for IoT security attack detection, combining CNN, LSTM, and attention mechanisms with a genetic algorithm-based optimization process. While not AD-specific, this architecture demonstrates the effectiveness of multi-component DL systems for achieving high detection accuracy with low false alarm rates across diverse attack categories - a design pattern directly transferable to AD event log analysis.

The CCT Dublin thesis [14] addresses insider threat detection through a hybrid deep learning approach that integrates LSTM-based sequence modeling with behavioral profiling, demonstrating that hybrid DL architectures achieve superior performance compared to single-model approaches for detecting subtle, behaviorally consistent insider threats in enterprise log data. This finding is particularly relevant for AD security, where insider threats represent one of the most difficult detection challenges due to their use of legitimate credentials and access patterns.

Neural-symbolic approaches represent an emerging direction for AD security. Opanovych [15] proposes enhancing rule-based threat detection with deep learning in AD environments, arguing that the combination of interpretable symbolic rules (encoding domain security expertise) with learned DL representations (capturing statistical patterns in event data) achieves both high detection accuracy and the explainability required for SOC deployment. This neural-symbolic paradigm addresses one of the most significant limitations of pure DL approaches: the opacity of detection decisions that undermines analyst trust.

3.5 Hybrid and Ensemble Model Architectures

Hybrid architectures that combine complementary learning paradigms have emerged as the most promising direction for comprehensive AD threat detection, driven by the empirical finding - documented in multiple reviewed works - that no single algorithm achieves adequate performance across the full spectrum of AD attack types.

RANSEC [16] proposes a hybrid ensemble learning framework specifically designed for ransomware detection in Active Directory, combining multiple base classifiers through ensemble voting to achieve robust detection of ransomware propagation patterns within AD-authenticated networks. The AD-specific design of RANSEC - incorporating knowledge of AD group policy, network share access patterns, and authentication event correlations - demonstrates the performance advantage of domain-tailored hybrid architectures over general-purpose IDS systems.

TSARKA SCIENCE
Volume 1 (2026), Issue 2
pp. 2–19
eISSN: 3134-6057
pISSN: 3134-6049

The enterprise access log anomaly detection framework [17] introduces a novel dataset and a hybrid ensemble approach combining multiple ML models to detect anomalies in enterprise access logs. The hybrid approach addresses the class imbalance problem inherent in enterprise log data - where malicious events are extremely rare - by combining an anomaly detection component (identifying statistically unusual events) with a classification component (categorizing identified anomalies into attack types), achieving superior performance compared to single-model baselines on the novel dataset.

The AI-powered AD system for predictive analytics and user behavior analysis [18] integrates behavioral modeling with ML-driven anomaly detection in an AD-native deployment architecture, demonstrating the practical feasibility of embedding ML-based threat prediction directly within AD infrastructure rather than relying on external SIEM platforms. This deployment approach reduces detection latency and eliminates the data transfer bottleneck that limits real-time detection performance in traditional SIEM-centric architectures.

Table 3. Hybrid and ensemble architectures for Active Directory threat detection.

Architecture	Target in AD	Key Innovation	Reference
1	2	3	4
Hybrid ensemble (RANSEC)	Ransomware in AD	AD-specific ensemble for ransomware	[16] RANSEC 2024
Hybrid ensemble + dataset	Enterprise access logs	Novel dataset + multi-model hybrid	[17] Ent. Logs 2025
AI behavioral + ML AD	User behavior analysis	Predictive AD-native deployment	[18] AI-AD 2025
Hybrid DL (CNN+LSTM+Attn)	IoT / general security	Self-optimized DL architecture	[13] Sagu 2022
Neural-symbolic	Rule-based + DL in AD	Explainable hybrid detection	[15] Opanovych 2025
Insider threat hybrid DL	Enterprise insider threat	LSTM + behavioral profiling	[14] CCT 2024

3.6 Graph-Based Approaches: Provenance Analysis and Attack Graphs

Graph-based methods represent the most architecturally sophisticated approach to AD threat detection, exploiting the inherent graph structure of AD environments - where users, machines, groups, and services form a complex interconnected net-

Volume 1 (2026), Issue 2
pp. 2–19

TSARKA SCIENCE

eISSN: 3134-6057
pISSN: 3134-6049

work of authentication and access relationships.

The HADES system [5] achieves detection of AD attacks - including Pass-the-Hash and Golden Ticket - through whole-system provenance analysis. HADES constructs provenance graphs from Windows Security Event Log telemetry, partitioning execution by logon session to manage graph complexity. The system is evaluated in real AD environments, demonstrating reliable detection of complex multi-stage attacks that evade signature-based defenses by combining individually benign actions into malicious attack sequences. This work represents the most rigorous empirical validation of provenance-graph-based AD attack detection published to date.

Attack graph-based approaches model the vulnerability landscape of AD environments as directed graphs of attack paths, enabling ML classification of AD configurations as vulnerable or secure. Nebbione and Calzarossa [12] demonstrate this approach on 220 generated AD environments, with Random Forest achieving F1-score of 0.91. Herranz-Oliveros et al. [6] extend the attack graph paradigm to unsupervised learning for lateral movement threat mitigation, applying clustering algorithms to identify high-risk lateral movement paths without requiring labeled attack data.

Neural network-based dynamic defense of AD attack graphs is addressed by multiple works. Chowdhury et al. [19] propose defending AD by combining neural network-based dynamic defense with evolutionary attack graph analysis, using a neural network to select optimal defensive actions in response to evolving attack graph states - a reinforcement learning-adjacent approach that adapts the defense strategy dynamically as attackers modify their techniques. The co-evolutionary GNN approach of [20] extends this framework, using GNN-approximated dynamic programming for co-evolutionary defense of AD attack graphs, achieving scalable defense optimization that accounts for the attacker's adaptive response to defensive countermeasures.

Deep generative models are applied to AD security by [21], which proposes extending AD graphs with honeypot users generated by deep generative models. This approach creates synthetic decoy user accounts that appear legitimate to automated attackers but trigger alerts when accessed, providing a proactive detection mechanism complementary to the reactive anomaly detection approaches that dominate the literature. The practical fixed-parameter algorithm framework of [22] addresses the computational complexity of AD attack graph defense optimization, providing polynomial-time approximation algorithms applicable to large enterprise AD environments.

3.7 SIEM, XDR Integration, and Operational Deployment

The operational deployment context for ML-based AD threat detection is the enterprise Security Information and Event Management (SIEM) platform or, increasingly, the Extended Detection and Response (XDR) platform that integrates telemetry from

TSARKA SCIENCE

Volume 1 (2026), Issue 2
pp. 2–19

eISSN: 3134-6057
pISSN: 3134-6049

endpoints, identity infrastructure, and network layers.

The Open XDR approach to AD security is addressed by two complementary works. The IEEE paper on Active Directory Open XDR cyber security techniques [23] proposes AD-specific detection algorithms integrated within an XDR architecture, achieving anomaly detection across multiple telemetry sources simultaneously - a critical capability for detecting sophisticated attacks that deliberately spread their footprint across multiple data streams to avoid single-source detection. Preprints.org [24] provides a systematic analysis of AI/ML integration in cybersecurity through Open XDR technology specifically applied to Active Directory, arguing that the XDR integration model enables more comprehensive behavioral context for ML models than SIEM-centric approaches.

The AI-powered AD system for predictive analytics [18] and the cyber security system for AD using AI-powered threat detection [25] both address the challenge of embedding ML inference directly within AD infrastructure to minimize detection latency. The Azure AD hybrid join optimization framework [26] demonstrates that ML can be applied to optimizing AD operational performance - including authentication latency and group policy propagation - as well as security detection, suggesting that ML-based AD management systems can serve dual security and operational purposes.

Automated cyber threat identification and profiling using ML [27] provides a framework for automated threat intelligence generation applicable to AD security operations, reducing the manual analyst effort required to classify and triage detected anomalies. ChatNVD [28] explores LLM-based vulnerability management for cybersecurity, a capability with direct application to AD security assessment - automated generation of security advisories and remediation recommendations for detected AD misconfigurations and vulnerabilities.

3.8 Datasets and Experimental Evaluation

The availability and quality of evaluation datasets is a critical determinant of research validity in AD threat detection. Table 4 presents the datasets used or proposed in the reviewed works.

Table 4. Datasets used in reviewed works for Active Directory and hybrid ML threat detection research.

Dataset / Environment	Source / Reference	Type	AD-Specific
1	2	3	4
Real AD environment (hundreds of users)	[3] Kotlaba 2021	Real AD Event 4769 logs (6 wks)	Yes

TSARKA SCIENCE

eISSN: 3134-6057
pISSN: 3134-6049

1	2	3	4
Simulated AD (attack/normal)	[1] Opanovych 2025	Simulated AD with APT attacks	Yes
Semi-synthetic AD logs (LM)	[8] Uppstromer 2019	PtH, PtT, AD enum. logs	Yes
220 generated AD environments	[12] Nebbione 2023	Artificially generated AD graphs	Yes
AD attack graph environment	[6] Herranz-Oliveros 2024	Lateral movement attack graphs	Yes
Enterprise access logs + novel	[17] Ent. Logs 2025	Hybrid dataset (novel proposed)	Partial
AD event logs + MITRE tactics	[9] ARKAIV 2025	Low-level logs + tactic mapping	Yes
CERT Insider Threat (AD LDAP)	[11] Haq 2022	Insider threat, NLP features	Partial
IoT security benchmarks	[13] Sagu 2022	NSL-KDD, CICIDS (comparison)	No

A critical and persistent gap in the reviewed literature is the absence of a large-scale, publicly available, labeled dataset of Windows Security Event Logs covering the full spectrum of AD authentication-layer attacks (Kerberoasting, PtH, Golden Ticket, DCSync). The majority of AD-specific results are obtained on either simulated environments [1], proprietary real-world logs [3], or semi-synthetic datasets [8] - none of which are publicly available for independent replication. This limits the reproducibility of reported results and the ability to perform meaningful cross-study performance comparison. It is important to clarify a potential ambiguity arising from Table 4: the 220 artificially generated environments reported by Nebbione and Calzarossa [12] are programmatically generated Active Directory configuration graphs used for AI-assisted security assessment - that is, they represent synthetic AD topology structures (sets of users, groups, machines, and their relationships) evaluated for the presence of attack paths, not labeled Windows Security Event Log traces of attack executions. These generated configurations are internal to the methodology of [12] and are not distributed as a public resource. Accordingly, they do not constitute a labeled Security Event Log dataset and do not contradict the stated absence of such a resource. The distinction is between AD configuration graphs (what [12] generates) and AD event log datasets capturing authentication-layer attack sequences (what the field lacks).

TSARKA SCIENCE
Volume 1 (2026), Issue 2
pp. 2–19
eISSN: 3134-6057
pISSN: 3134-6049

3.9 Identified Research Gaps and Open Challenges

Synthesis of the reviewed literature reveals five primary research gaps that collectively define the research agenda for the proposed dissertation work:

First, the absence of a unified hybrid model validated on AD authentication event data. While hybrid architectures combining CNN, LSTM, GNN, and ensemble methods have demonstrated strong performance in related domains, no reviewed work proposes and validates a comprehensive hybrid framework specifically designed for Windows Security Event Log analysis across the full spectrum of AD attack types documented in Table 1. Opanovych [1] identifies the need for combined approaches but does not propose a unified architecture; HADES [5] achieves provenance-graph detection but does not incorporate statistical ML components.

Second, the lack of multi-stage kill-chain detection capability. Real-world APT campaigns follow multi-stage sequences in which individual stages appear benign in isolation. Matsuda et al. [2] and Opanovych [1] demonstrate detection of individual techniques, but neither integrates a kill-chain model that contextualizes individual anomalies within the broader attack narrative. ARKAIV [9] represents the closest work to this direction, modeling tactic sequences, but focuses on exfiltration prediction rather than comprehensive kill-chain detection.

Third, the scarcity of labeled, publicly available AD attack datasets. As documented in Section 3.8, no reviewed work introduces a public large-scale dataset of Windows Security Event Logs with ground-truth labels for authentication-layer AD attacks. The enterprise access log anomaly detection paper [17] proposes a novel dataset, but it focuses on general enterprise access logs rather than AD-specific Security Event Logs.

Fourth, unresolved real-time deployment challenges. The HADES system [5] is evaluated in real AD environments but does not report quantitative latency benchmarks at enterprise domain controller scale. The AI-powered AD systems [18, 25] report operational deployment but do not provide systematic performance evaluation across diverse attack scenarios. No reviewed work establishes a comprehensive benchmark for real-time ML inference performance at production AD event rates.

Fifth, limited application of explainable AI (XAI) techniques. The neural-symbolic approach of Opanovych [15] provides partial interpretability through symbolic rule integration, and Nebbione and Calzarossa [12] provide attack path visualization. However, systematic application of SHAP, LIME, or attention-based interpretability to hybrid AD threat detection models - enabling SOC analysts to understand detection rationale in terms of MITRE ATT&CK technique mappings - has not been demonstrated in the reviewed literature.

Volume 1 (2026), Issue 2
pp. 2–19

TSARKA SCIENCE

eISSN: 3134-6057
pISSN: 3134-6049

Table 5. Summary of identified research gaps with supporting evidence and proposed research directions.

Research Gap	Evidence from Review	Proposed Direction
1	2	3
No AD-specific hybrid model	Opanovych [1]: need combined approaches; HADES [5]: no ML integration	Hybrid CNN-LSTM-RF on AD Event Logs
No kill-chain detection	Matsuda [2], Opanovych [1]: single-technique	Attack intent sequence modeling
No public AD attack dataset	All AD results on private/simulated data	Labeled dataset generation or collection
Real-time at DC scale	HADES [5]: no latency benchmarks	Efficient architecture + latency evaluation
Limited XAI in AD detection	Nebbione [12]: paths only; Opanovych [15]: partial	SHAP + ATT&CK narrative generation

4. Conclusion

This literature review has systematically examined 30 peer-reviewed and technically validated works on machine learning-based threat detection in Active Directory environments, synthesizing findings across attack taxonomies, feature engineering, classical ML, deep learning, hybrid architectures, graph-based methods, and operational deployment contexts.

The principal finding is a convergence of evidence across the reviewed corpus that no single algorithmic paradigm is sufficient for comprehensive AD threat detection. Kotlaba et al. [3] demonstrate the limitations of threshold-based rules and the superiority of One-Class SVM for Kerberoasting detection; Opanovych [1] establishes empirically that LSTM, Autoencoder, and graph-based algorithms each have distinct strengths across different AD attack categories; Nebbione and Calzarossa [12] validate graph-ML integration for AD security assessment; and works on hybrid ensemble architectures [16, 17] consistently outperform single-model baselines. This convergence establishes the hybrid multi-paradigm architecture as the most promising direction for the proposed research.

The most significant technical contributions identified in the reviewed literature are: (1) the provenance-graph-based approach of HADES [5] for detecting complex multi-stage AD attacks in real environments; (2) the tactic-sequence prediction frame-

TSARKA SCIENCE

Volume 1 (2026), Issue 2
pp. 2–19

eISSN: 3134-6057
pISSN: 3134-6049

work of ARKAIV [9] for bridging the gap between low-level event logs and high-level MITRE ATT&CK attack modeling; (3) the neural-symbolic integration proposed by Opanovych [15] for achieving both high detection accuracy and SOC-interpretable explanations; and (4) the attack-graph-based ML assessment framework of Nebbione and Calzarossa [12] for proactive AD vulnerability identification.

Five research gaps have been identified - the absence of a unified AD-specific hybrid model, the lack of kill-chain-aware detection, the scarcity of labeled public AD datasets, unresolved real-time deployment challenges, and limited XAI application - that collectively define the scope and contribution potential of the proposed dissertation research. Addressing these gaps through the development, evaluation, and deployment of a novel hybrid machine learning framework for threat prediction and detection in Active Directory environments based on systematic Windows Security Event Log analysis represents a high-impact research direction with clear practical value for enterprise cybersecurity. The envisioned architecture of the proposed framework integrates three complementary processing layers: (1) a feature extraction layer combining statistical aggregation of per-user authentication event sequences with MITRE ATT&CK tactic-level encoding of Windows Security Event Log streams; (2) a hybrid detection layer comprising a BiLSTM module for sequential anomaly detection in authentication event sequences, a Graph Neural Network (GNN) module operating on dynamically constructed provenance graphs of AD session activity, and a Random Forest / XGBoost ensemble classifier for high-confidence attack-type attribution; and (3) an explainability layer applying SHAP-based feature attribution to map detection decisions to interpretable MITRE ATT&CK technique identifiers for SOC analyst consumption. The three detection components are combined through a late-fusion aggregation mechanism that produces a unified risk score and attack-type prediction. Regarding the dataset gap identified as a primary research challenge, the authors plan to generate, label, and publicly release a structured dataset of Windows Security Event Logs collected from a controlled Active Directory laboratory environment encompassing the full spectrum of authentication-layer attacks documented in Table 1. Publication of this dataset as an open research resource is explicitly identified as one of the primary contributions of the planned dissertation work, with the intention of enabling reproducible benchmarking across future AD threat detection research.

References

1. Opanovych M.Y. Anomaly Detection in Active Directory for APT Attack Detection: Methods, Effectiveness, and Limitations // Cybersecurity: Education, Science, Technique. - 2025. - Vol. 2(30). - P. 9–15. DOI: 10.28925/2663-4023.2025.30.915.
2. Matsuda W., Fujimoto M., Mitsunaga T. Detecting APT Attacks Against Active Directory Using Machine Learning // Proc. IEEE Conference on Application, In-

formation and Network Security (AINS 2018). - 2018. - P. 60–65. DOI: 10.1109/AINS.2018.8631486.

3. Kotlaba L., Buchovecka S., Lorencz R. Active Directory Kerberoasting Attack: Detection using Machine Learning Techniques // Proc. 7th International Conference on Information Systems Security and Privacy (ICISSP 2021). - 2021. - P. 376–383. DOI: 10.5220/0010202803760383.
4. Mabika M. Supervised Learning on Active Directory with Overcoming Cybersecurity Challenges. Master's thesis, National College of Ireland. - 2024. - URL: https://norma.ncirl.ie/8222/
5. Holder C., Goel A. Detecting and Investigating Active Directory Attacks via Whole-System Provenance Analysis // IEEE Access. - 2025. - DOI: 10.1109/ACCESS.2025.11175589. [Also: arXiv:2407.18858].
6. Herranz-Oliveros D., Tejedor-Romero M., Gimenez-Guzman J.M., De La Cruz-Piris L. Unsupervised Learning for Lateral-Movement-Based Threat Mitigation in Active Directory Attack Graphs // Electronics. - 2024. - Vol. 13, No. 19. - P. 3944. DOI: 10.3390/electronics13193944.
7. Uzun A.E. An Analysis of Kerberoasting Attack and Detection with Supervised Machine Learning Algorithms. Master's thesis, Middle East Technical University. - 2023. - URL: https://open.metu.edu.tr/handle/11511/112647
8. Uppstromer V., Raberg H. Detecting Lateral Movement in Microsoft Active Directory Log Files: A Supervised Machine Learning Approach. Master's thesis, Blekinge Institute of Technology. - 2019. - URL: https://www.diva-portal.org/smash/get/diva2:1333721/FULLTEXT01.pdf
9. Hakim A.R., Ramli K., Salman M., Pranggono B., Agustina E.R. ARKAIV: Predicting Data Exfiltration Using Supervised Machine Learning in Active Directory Environments // IEEE Access. - 2025. - DOI: 10.1109/ACCESS.2024.10818683.
10. Automated Detection of Ransomware in Windows Active Directory Domain. Preprint. - 2024. - URL: https://d197for5662m48.cloudfront.net/documents/publicationstatus/225955/preprint_pdf/1ca9bb504df1c0d1d47524910f563602.pdf
11. Haq M.A., Khan M.A.R., Alshehri M. Insider Threat Detection Based on NLP Word Embedding and Machine Learning in Active Directory Logs // Intelligent Automation & Soft Computing. - 2022. - Vol. 33, No. 1. - P. 235–252. DOI: 10.32604/iasc.2022.021430.
12. Nebbione G., Calzarossa M.C. A Methodological Framework for AI-Assisted Security Assessments of Active Directory Environments // IEEE Access. - 2023. - Vol. 11. DOI: 10.1109/ACCESS.2023.3244490.
13. Sagu A., Gill N.S., Gulia P., Chatterjee J.M., Priyadarshini I. A Hybrid Deep Learning Model with Self-Improved Optimization Algorithm for Detection of Security Attacks in IoT Environment // Future Internet. - 2022. - Vol. 14, No. 10. - P. 301. DOI: 10.3390/fi14100301.

TSARKA SCIENCE

Volume 1 (2026), Issue 2
pp. 2–19

eISSN: 3134-6057
pISSN: 3134-6049

14. Enhancing Insider Threat Detection Through A Hybrid Deep Learning Approach. Technical Report. - CCT Dublin, 2024. - URL: https://arc.cct.ie/cgi/viewcontent.cgi?article=1133&context=ict
15. Opanovych M.Y. Neural-Symbolic Approaches to Active Directory Security: Enhancing Rule-Based Threat Detection with Deep Learning. ResearchGate Preprint. - 2025. - URL: https://www.researchgate.net/publication/399235150
16. RANSEC: Hybrid Ensemble Learning-based Secure Approach for Ransomware Detection in Active Directory // Journal of Applied Security and Technology Trends (JASTT). - 2024. - URL: https://jastt.org/index.php/jasttpath/article/view/555
17. A Novel Dataset and a Hybrid Ensemble Approach for Anomaly Detection in Enterprise-Access-Logs. ResearchGate Preprint. - 2025. - URL: https://www.researchgate.net/publication/392210704
18. Development of an AI-Powered Active Directory System for Predictive Analytics and User Behavior Analysis // IEEE Xplore. - 2025. - DOI: 10.1109/access.2025.11376946.
19. Chowdhury A., et al. Defending Active Directory by Combining Neural Network Based Dynamic Defense with Evolutionary Attack Graph // arXiv:2204.03397. - 2022.
20. Co-Evolutionary Defence of Active Directory Attack Graphs via GNN-Approximated Dynamic Programming // arXiv:2505.11710. - 2025.
21. Deep Generative Models to Extend Active Directory Graphs with Honeypot Users // arXiv:2109.06180. - 2021.
22. Practical Fixed-Parameter Algorithms for Defending Active Directory Attack Graphs // arXiv:2112.13175. - 2021.
23. Active Directory Open XDR Cyber Security Techniques to Detect Anomalies // IEEE Xplore. - 2025. - DOI: 10.1109/access.2025.11081224.
24. Integrating AI/ML in Cybersecurity: An Analysis of Open XDR Technology and its Application in Active Directory. Preprints.org. - 2023. - DOI: 10.20944/preprints202312.0205.v1.
25. Development of a Cyber Security System for Active Directory Using AI-Powered Threat Detection and Response // IEEE Xplore. - 2025. - DOI: 10.1109/access.2025.11376826.
26. Performance Optimization of Hybrid Azure AD Join Across Environments Using Machine Learning // JISEM Journal. - 2025. - URL: https://jisem-journal.com/index.php/journal/article/view/8897
27. Automated Emerging Cyber Threat Identification and Profiling Using Machine Learning // IEEE Access. - 2023. - DOI: 10.1109/ACCESS.2023.10077593.
28. ChatNVD: Advancing Cybersecurity Vulnerability Management with LLMs and

Hybrid Models // IEEE Access. - 2025. - DOI: 10.1109/ACCESS.2025.11369932.

29. Optimizing Cyber Defense in Dynamic Active Directories // arXiv:2406.19596. - 2024.

30. MITRE Corporation. ATT&CK for Enterprise v14. - 2024. - URL: https://attack.mitre.org.

Information about authors

Tastemirov Azamat Azizbekovich –
Higher Education, TSARKA GROUP, Astana, Kazakhstan
e-mail: tazamat@cybersec.kz
ORCID: https://orcid.org/0009-0009-6215-3081

Bekbolatov Miraskhan Bekbolatuly –
Higher Education, TSARKA GROUP, Astana, Kazakhstan
e-mail: m.bekbolatov@cybersec.kz
ORCID: https://orcid.org/0009-0006-9617-0048