

Hybrid AI framework for anomaly detection and root cause analysis in multi-agent systems

Tahri Rachid¹, Ouammou Abdellah¹, Lasbahani Abdellatif², Abdessamad Jarrar³, Balouki Youssef⁴

¹Faculty of Sciences and Technologies, Hassan First University, Settat, Morocco

²Faculty of Sciences and Technologies, Sultan Moulay Slimane University, Beni Mellal, Morocco

³Faculty of Sciences, Mohammed First University, Oujda, Morocco

⁴Higher School of Technology, Mohammed V University, Salé, Morocco

Article Info

Article history:

Received Oct 28, 2023

Revised Oct 19, 2025

Accepted Nov 10, 2025

Keywords:

Anomaly detection
Artificial intelligence
Knowledge graphs
Machine learning
Root cause analysis

ABSTRACT

Anomaly detection and root cause analysis (RCA) are critical for securing intelligent systems against evolving threats. Traditional models often suffer from high false alarms, weak adaptability to streaming contexts, and limited interpretability. This work proposes a hybrid artificial intelligence (AI) framework that integrates machine learning (ML) with prior knowledge, semantic rules, and bio-inspired modeling. The approach strengthens detection of diverse attacks, including DoS/DDoS, Probe, U2R, and R2L, while reducing human intervention. Experiments on the NSL-KDD dataset demonstrate that our method decreases spurious alerts by up to 90%, improves accuracy by 2–4%, and reduces false positives/negatives by about 4%. Beyond statistical gains, the framework ensures robustness in real-time environments, offering interpretable and scalable anomaly detection for heterogeneous systems. These results highlight the potential of hybrid symbolic–subsymbolic AI to enhance reliability in next-generation security infrastructures.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Tahri Rachid

Faculty of Sciences and Technologies, Hassan First University

Km 3, B.P.: 577 Route de Casablanca, Settat, Morocco

Email: rachid.tahri.tr@gmail.com

1. INTRODUCTION

The rapid expansion of artificial intelligence (AI) and the internet of things (IoT) has led to massive streams of heterogeneous data reflecting device activity, environmental conditions, and contextual dynamics. Sensors, which now permeate domains such as healthcare, logistics, and industrial automation, produce continuous flows of information on physical units and their interactions. While these advances create unprecedented opportunities for intelligent services, they also raise major challenges: how to integrate devices in a secure and adaptive way, how to analyze data efficiently, and how to detect abnormal behaviors with minimal human intervention.

From a security perspective, anomaly detection and root cause analysis (RCA) remain central techniques. Anomaly detection highlights events deviating from expected patterns, revealing incidents such as technical failures or fraudulent activities [1], while RCA (or deep cause analysis (DCA)) investigates underlying causes to support corrective actions [2]. Together, anomaly detection and DCA form the foundation for resilient intelligent systems. Figure 1 illustrates the general workflow: data from multiple sources undergo preprocessing, models detect deviations from normal behavior, and corrective mechanisms are triggered when

anomalies are confirmed. For example, in wildfire monitoring, deviations in temperature or humidity sensors can activate preventive interventions.

Despite their utility, existing approaches still face key limitations. Conventional anomaly detection models trained on static data often fail in dynamic environments, where evolving conditions—such as climate shifts or sensor drift—are misclassified as anomalies. Likewise, faulty devices may produce false alarms, requiring costly human intervention [3]. Moreover, the majority of current systems lack contextual awareness, relying solely on raw data without exploiting prior knowledge. This leads to excessive false positives and limits adaptability in real-time applications. Recent advances in AI, such as graph neural networks (GNNs), transformer-based time-series models, and generative adversarial network (GAN)-based anomaly detection, highlight the potential of hybrid symbolic–subsymbolic approaches, yet their integration into RCA for multi-agent environments remains underexplored. The goal of this study is to reinforce anomaly detection by combining machine learning (ML) with semantic knowledge and contextual reasoning, thereby improving accuracy, adaptability, and automation. Specifically, our contributions are as follows: i) reducing false alarms and fraudulent alerts by integrating structured prior knowledge into the detection pipeline; ii) enhancing detection and corrective actions across dynamic and heterogeneous environments [4]; iii) improving interpretability by automating root cause identification through semantic rules; and iv) enabling adaptive detection behavior under context changes, reducing reliance on human experts.

The remainder of this paper is structured as follows. Section 2 reviews related work and identifies existing gaps. Section 3 presents our proposed hybrid framework, combining AI models with semantic reasoning. Section 4 details the algorithms used, followed by section 5 reporting experimental evaluation. Finally, section 6 discusses findings, and section 7 concludes with future directions.

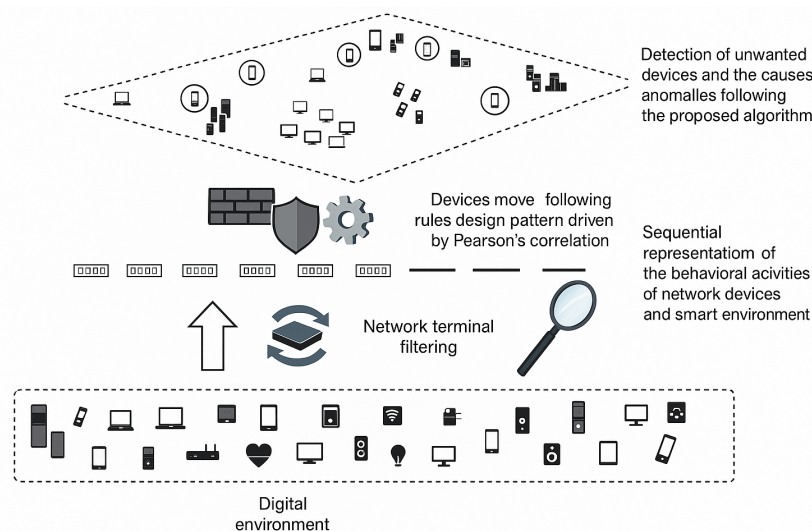


Figure 1. Anomaly detection process

2. RELATED WORK

Research on anomaly detection and RCA has evolved along multiple complementary lines. Early studies mainly relied on raw-data-driven learning models, while more recent contributions emphasize the use of contextual knowledge and prior information to enhance robustness and interpretability. Several works target anomaly detection in unstructured datasets without predefined models, thresholds, or rules [5], [6]. Today, most approaches rely on ML, in both supervised and unsupervised modes, depending on the availability of labels. While supervised models achieve solid results, unsupervised detection remains challenging due to the unpredictable variability of real-world data. Statistical and tree-based methods have also been explored to characterize normal traffic and detect deviations [7]–[10].

The semantic web (SW) has been investigated to represent knowledge with RDF and ontologies [11]. Ontologies enable reasoning over domain entities, but integration with ML often requires embeddings or graph-based transformations. Approaches such as relational graph convolutional networks (RGCN) and

inductive logic programming (ILP) attempt to bridge this gap [12]-[16], though explicit use of SW concepts for anomaly detection remains limited.

Other research explores metaheuristics and multi-agent paradigms. For instance, IoT2Vec models device/service footprints in IoT networks [17], while bio-inspired agents have been applied to streaming anomaly detection [18]. These solutions are adaptive and decentralized but often lack context-awareness regarding sensor flows or environmental changes. Policy-based approaches offer greater explainability but require significant human involvement and long development cycles [2].

Domain-specific contributions also exist. For credit card fraud detection, Mniai and Jebbari [19] combined support vector data description (SVDD) with particle swarm optimization (PSO), while in network security, Song *et al.* [20] proposed an Anti-DoS Duplicate Address Detection model. Nonetheless, scalability, adaptability, and interoperability remain open challenges. Table 1 summarizes these families of approaches, highlighting strengths and limitations.

In summary, the literature highlights persistent challenges: reliance on raw data without prior knowledge, high false alarms, poor adaptation to dynamic streams, RCA designs still dependent on experts, limited interpretability, and lack of self-adaptation or interoperability. Unlike these prior works, our contribution proposes a hybrid framework that integrates ML, semantic knowledge, and bio-inspired multi-agent modeling. This design explicitly addresses the issues of adaptability, interpretability, and automation, paving the way for more robust and context-aware anomaly detection in intelligent systems.

Table 1. Comparative summary of anomaly detection approaches

Category	Representative methods/ works	Strengths	Limitations
ML-only models	Naïve Bayes, decision trees, support vector machine (SVM), random forest (RF) [5]-[10], [19], [20]	Simple, scalable, reasonable accuracy	High false positives, weak adaptability, limited interpretability
Semantic/ontology-based	RDF/OWL, SW ontologies [21]-[12], [13], [15], [16]	Encodes domain knowledge, improves explainability	Limited ML integration, costly vectorization, rarely applied to anomaly detection
Hybrid ML + knowledge	RGCN, ILP, neuro-symbolic AI [14], [15]	Context-aware, reasoning, generalization	Few implementations, scalability issues, needs expert rules
Metaheuristic/multi-agent	IoT2Vec, bio-inspired agents [17], [18]	Adaptive, decentralized, suitable for IoT	Limited contextual awareness, lack of predictive evaluation, development overhead
Recent AI-based methods	GAN-based detectors, GNNs, transformer models (DeepAR, TFT, Informer) [22]	Capture complex patterns, handle streams, reduce false positives	Few applied to RCA, limited multi-agent benchmarks

3. PROPOSED APPROACH

Our approach combines bio-inspired multi-agent dynamics, semantic reasoning, and hybrid ML models to enhance anomaly detection and RCA. The design emphasizes adaptability, interpretability, and low human intervention.

3.1. Bio-inspired model

Swarm intelligence (SI) provides self-organization and adaptive coordination in distributed environments [23]-[25]. We adopt a flocking model [26], [27], where each agent occupies a 2D grid cell (x, y) and evolves according to the canonical rules of cohesion, separation, and alignment [28]. The resulting velocities are aggregated to update positions:

$$\vec{v}_a = \frac{1}{n} \sum_{i=1}^n \vec{v}_i, \quad \vec{u}_s = \sum_{i=1}^n \frac{\vec{v}_i + \vec{u}_{AC}}{d(fm_i, A_c)}, \quad \vec{u}_c = \frac{1}{n} \sum_{i=1}^n (pos_i - pos_c) \quad (1)$$

To refine discrimination, a similarity-driven term is added:

$$v_s = \sum_{i=1}^n S(fm_i, A_c) d(pos_i, pos_c), \quad v_d = \frac{1}{\sum_{i=1}^n S(fm_i, A_c) d(pos_i, pos_c)} \quad (2)$$

The final velocity is a weighted combination:

$$v_{AC}^F = w_a v_a + w_s v_s + w_c v_c + w_d v_d \quad (3)$$

This mechanism clusters coherent agents and separates outliers, improving anomaly detection in dynamic contexts.

3.2. Hybrid anomaly detection+deep cause analysis framework

The pipeline integrates ML with contextual rules to increase robustness. Main steps include: i) embedding prior knowledge into detection profiles; ii) filtering spurious alerts with semantic criteria; iii) balancing false positives/negatives; iv) automating RCA routines to reduce human workload. Figure 2 shows the extended workflow: data ingestion → preprocessing → detection → RCA actions, enriched with semantic knowledge.

3.2.1. Matrix representation and region selection

Knowledge graphs are encoded as matrices, where rows represent subject–object pairs and columns represent relation types. Abnormal subregions (RoIs) are selected by rule-based virtual agents. These agents focus detection on suspicious links (Figure 3).

3.2.2. Vector features and GAN bridge

Features are vectorized for computation (Figure 4). A GAN then reconstructs candidate graphs, ensuring semantic consistency with the original knowledge base (Figure 5). This preserves interpretability while improving generalization.

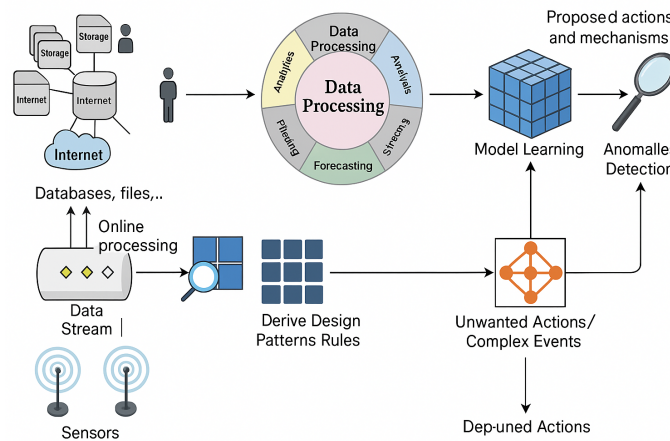


Figure 2. Extended anomaly detection+deep cause analysis (AD+DCA) workflow with semantic and ML integration

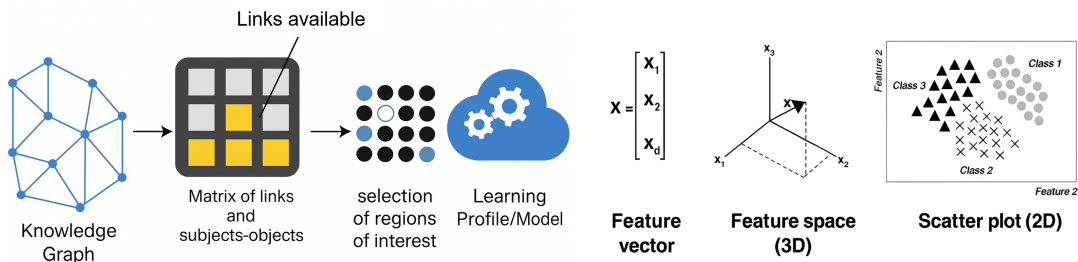


Figure 3. Matrix-based knowledge representation with rule-driven RoI selection

Figure 4. Feature vector representation used for detection and RCA

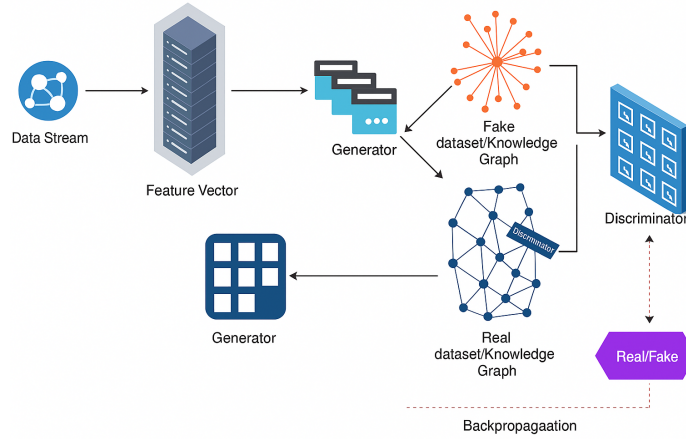


Figure 5. GAN-based reconstruction for interpretable anomaly analysis

3.2.3. Streaming patterns and context likelihood

Streaming data induces drift. We extract rules as *design patterns*, updating profiles in real time. The likelihood that a device is anomalous under a context is computed as:

$$P(\text{context}) = \frac{\text{similarity}(\text{devices}_i \mid \text{context})}{\sum_k \text{similarity}(\text{devices}_k \mid \text{context})} \quad (4)$$

with Pearson correlation used as similarity, effective with Device2Vec embeddings [22].

3.3. Algorithm

Algorithm 1 summarizes the detection of anomaly causes. Each device is mapped to an agent. If contextual similarity falls below a threshold, corrective rules are applied.

Algorithm 1 Pseudocode for detecting causes of anomalies

```

1: while true do
2:   Construction(), Mapping(), Random.distribution()
3:   for i ← 1 to Iterations do
4:     for j ← 1 to Na do
5:       if agentj.context.values() ≠ Vj then
6:         c ← agentj.complexity()
7:         for k ← 1 to Na,j do
8:           for z ← 1 to Na,j,z do
9:             if Pearson(va,j, va,k, va,z) ≥ ε then
10:              Apply.all()
11:            else
12:              Separate()
13:            end if
14:          end for
15:          Compute.C(), Compute.Velocity(), Move()
16:        end for
17:      end if
18:    end for
19:    Update.Rules()
20:  end for
21: end while

```

3.4. Attack taxonomy (NSL-KDD)

Our dataset comprises a diverse set of attack types relevant to modern network security contexts. Specifically, we identified 11 sub-types within the DoS/DDoS class, 6 under the Probe class, 7 in the User to Root (U2R) class, and 15 related to Remote to Local (R2L) attacks. This categorization provides a comprehensive foundation for analyzing the prevalence and behavior of different cyberattack vectors.

A detailed analysis of the NSL-KDD dataset reveals the statistical distribution of attack classes, as illustrated in Figure 6. Normal traffic constitutes 53%, while DoS/DDoS accounts for 36%, Probe for 9%, U2R for 0.03%, and R2L for 0.78%. These figures offer insight into the relative frequency of each threat, guiding the focus of our detection strategies.

Figure 6(a) illustrates the distribution of DoS/DDoS attacks across the dataset. This visualization helps in identifying which attack types are most prevalent and require closer attention. Understanding these distributions is essential to prioritize detection mechanisms. Figure 6(b) presents the distribution of Probe-type attacks recorded in our dataset. These attacks are typically used to gather information about a target network, often preceding more severe intrusions. Understanding their frequency and behavior is essential for implementing proactive detection and prevention mechanisms. Figure 6(c) illustrates the distribution of user to root (U2R) attacks. These attacks occur when a user with limited access privileges attempts to gain root-level control over a system. Although less frequent, they pose a critical threat due to the high-level access they can grant to malicious actors.

Figure 6(d) displays the occurrence of remote to local (R2L) attacks within the analyzed dataset. These attacks aim to exploit vulnerabilities in remote systems to gain unauthorized access at the local level. Despite their relatively low frequency, they remain significant due to their potential for unnoticed infiltration. Figure 6(e) provides an overview of the general distribution of all attack classes included in the dataset. This summary highlights the predominance of DoS/DDoS attacks, which account for the majority of cases. Such visualization helps identify which categories require the most attention in terms of detection and mitigation strategies. As evident from the pie chart, DoS/DDoS attacks exhibit the highest rate at 36% in our dataset, while U2R and R2L attacks are notably scarce. This observation indicates a direct proportionality between our dataset and contemporary internet traffic attacks. For our study, we will designate `attack_class = 1`, specifically targeting DoS/DDoS attacks, as the dependent variable. Figure 6(f) focuses on DDoS attacks segmented by protocol type, specifically TCP, UDP, and ICMP. Each protocol exhibits unique vulnerabilities that attackers exploit to disrupt network services. Understanding the distribution by protocol allows security systems to fine-tune their monitoring and mitigation mechanisms.

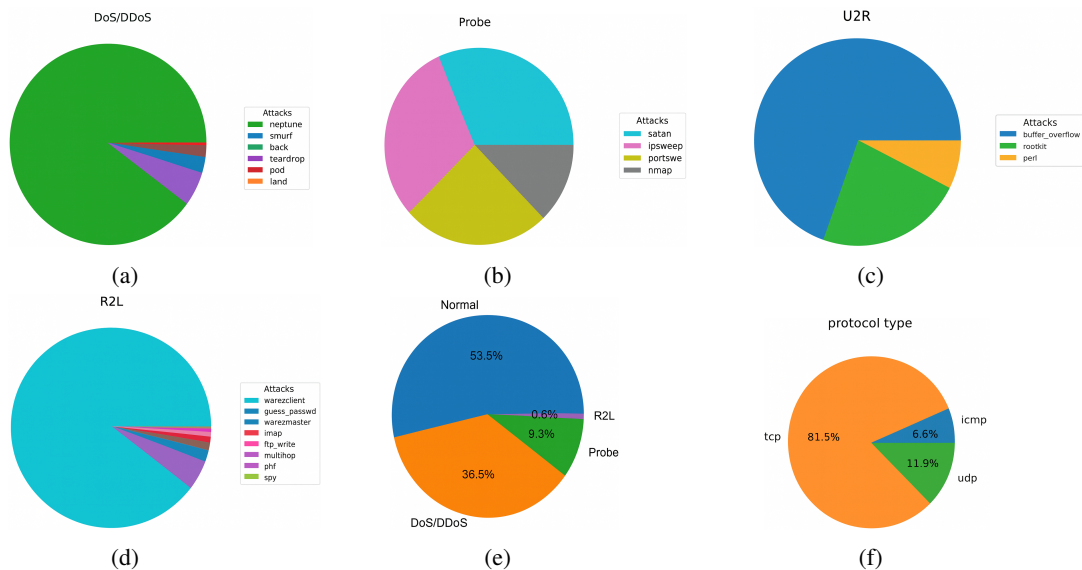


Figure 6. NSL-KDD class distributions and protocol breakdown: (a) DDoS, (b) Probe, (c) U2R, (d) R2L, (e) classes, and (f) by protocol

3.5. Comparative positioning

Table 2 provides a comprehensive comparison of various anomaly detection methods, highlighting their main objectives, strengths, and limitations. It helps to contextualize our proposed framework within the existing landscape by illustrating how it addresses specific gaps. This positioning underscores the unique contributions and advantages of our approach relative to other established techniques.

Table 2. Comparison of anomaly detection approaches

Approach	Strengths	Limitations	Gap Addressed by Our Work
ML-only models	Good accuracy with labeled data	High false alarms, poor adaptation	Context-aware filtering, false-alarm reduction
Semantic web	Reasoning via ontologies	Weak real-time performance, integration cost	Lightweight semantic rules in streaming pipeline
Metaheuristic/ Multi-agent	Adaptivity, decentralization	Limited context awareness	Bio-inspired clustering with similarity weighting
Hybrid (e.g., SVDD+PSO)	Better detection in specific domains	Scalability and efficiency issues	Generalizable across domains, stream-ready
Proposed model	High accuracy, interpretability, real-time RCA	Requires integration of heterogeneous components	Demonstrates synergy of ML + semantics + swarm dynamics

4. ALGORITHMS USED IN THE MANIPULATION OF OUR DATASET

This section outlines the main algorithms applied to the NSL-KDD dataset for anomaly detection and cause analysis.

4.1. Naïve Bayes

Naïve Bayes (NB) is a probabilistic classifier based on Bayes' theorem under the assumption of conditional independence among features. For class y and input vector $x = (x_1, \dots, x_n)$:

$$P(y | x) \propto P(y) \prod_{i=1}^n P(x_i | y) \quad (5)$$

We use the Gaussian variant (GaussianNB), where each feature likelihood is modeled by a normal distribution:

$$P(x_i | y) = \frac{1}{\sqrt{2\pi\sigma_{y,i}^2}} \exp\left(-\frac{(x_i - \mu_{y,i})^2}{2\sigma_{y,i}^2}\right) \quad (6)$$

with $\mu_{y,i}$ and $\sigma_{y,i}^2$ estimated by maximum likelihood. Prediction is obtained by maximizing:

$$\hat{y} = \arg \max_y \left\{ \log P(y) + \sum_{i=1}^n \log P(x_i | y) \right\} \quad (7)$$

Despite its simplicity, GaussianNB remains a robust baseline for high-dimensional tabular data.

4.2. Long short-term memory for sequential detection

Long short-term memory (LSTM) networks extend recurrent neural networks with memory cells and gating mechanisms to capture long-range dependencies. At each step t , the gates are defined as:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \quad i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (8)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad \tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (9)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \quad h_t = o_t \odot \tanh(C_t) \quad (10)$$

Parameters \mathbf{w} are optimized with gradient descent:

$$\mathbf{w} \leftarrow \mathbf{w} - \alpha \nabla_{\mathbf{w}} \mathcal{L} \quad (11)$$

We exploit LSTM to model traffic as sequences, enabling real-time anomaly detection under drifts or bursts. Figure 7 summarizes its architecture.

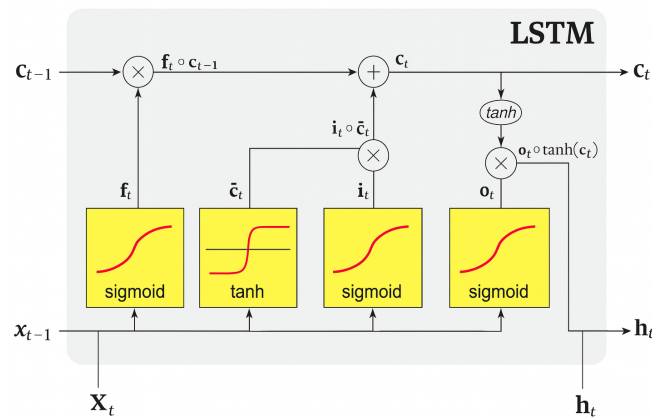


Figure 7. LSTM operations: gates (f_t, i_t, o_t), candidate state \tilde{C}_t , cell state C_t , and hidden output h_t

4.3. Deep cause analysis with semantic rules

To enhance interpretability, we complement ML predictions with a DCA layer. Detected anomalies are linked to semantic rules describing deviations from normative behavior. Rule execution provides causal paths and corrective actions (Figure 8).

4.4. Comparative summary of algorithms

Table 3 provides a detailed comparison of the three algorithms employed in this study, highlighting their specific objectives, strengths, and limitations. It illustrates how each algorithm contributes to anomaly detection and root cause analysis, emphasizing aspects such as their performance accuracy, speed, adaptability to streaming data, and interpretability. This summary helps in understanding the operational and technical trade-offs involved in implementing these methods in real-world scenarios. It also guides the selection of appropriate algorithms based on system constraints, desired outcomes, and scalability concerns.

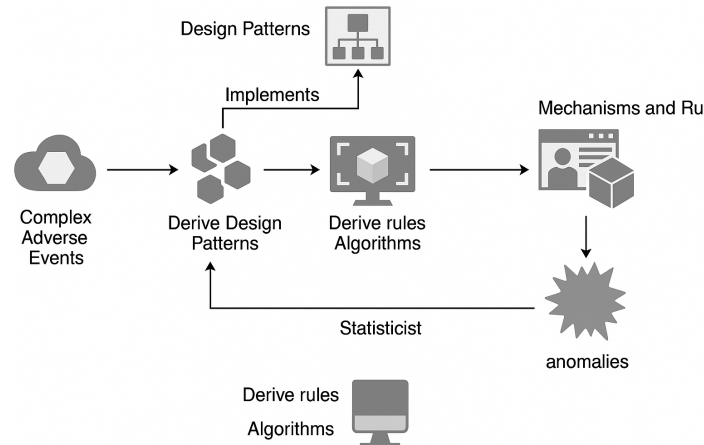


Figure 8. Rule-based DCA: anomalies are mapped to causal interpretations and remedial actions

Table 3. Summary of algorithms used in our dataset analysis

Algorithm	Objective	Strengths	Limitations
Naïve Bayes (GaussianNB)	Probabilistic classification of traffic records	Fast, scalable, robust to high dimensions, low computational cost	Assumes feature independence, limited handling of complex correlations
LSTM	Sequence modeling for real-time detection	Captures temporal dependencies, adapts to drifts/bursts, suitable for streaming	Computationally expensive, requires tuning, sensitive to long training times
DCA with semantic rules	Interpret anomaly causes using domain knowledge	Increases interpretability, provides causal paths and actionable insights	Rule construction may require expert input, scalability depends on rule base size

5. RESULTS

The evaluation of our framework was conducted on the NSL-KDD dataset. Three experimental stages were considered: i) baseline classifiers on raw features, ii) the same models with feature normalization, and iii) our extended AD+DCA process enriched with prior knowledge.

5.1. Baseline and normalized classifiers

In the first stage, GaussianNB achieved 80% accuracy (Figure 9), decision tree reached 88% (Figure 10), while k-nearest neighbors (KNN) performed worst at 73%. These outcomes confirm the limitations of models relying solely on unprocessed attributes. RF obtained the best performance at 90% (Figure 11), while SVM matched the decision tree baseline with 88%. Figure 12 illustrates system behavior during these simulations. With normalized features, GaussianNB slightly decreased to 75%, decision tree stayed at 88%, while KNN improved markedly to 87% (Figure 13). RF remained the most accurate at 90% (Figure 14), and SVM stayed stable at 88%. These results indicate that normalization particularly benefits distance-based models such as KNN.

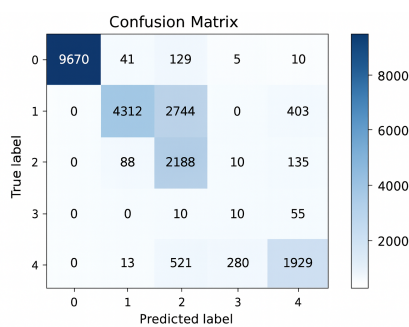


Figure 9. Confusion matrix (GaussianNB)

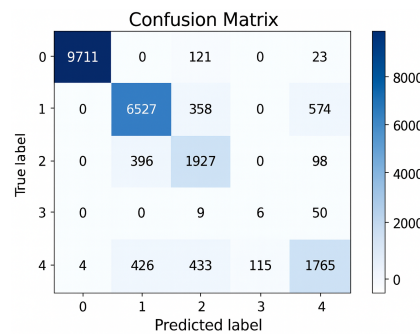


Figure 10. Confusion matrix (decision tree)

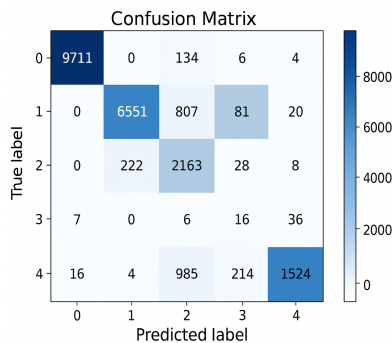


Figure 11. Confusion matrix (RF)

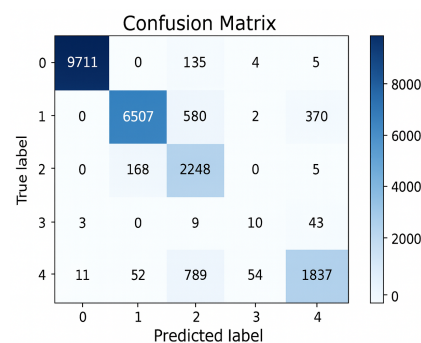


Figure 12. Simulation of normal vs. anomalous behaviors

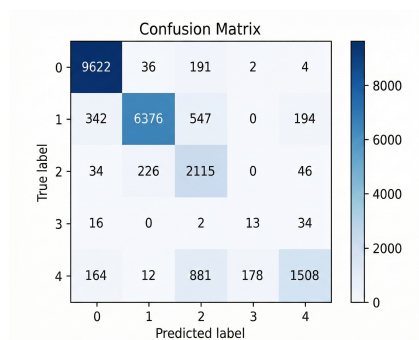


Figure 13. Confusion matrix (KNN with normalization)

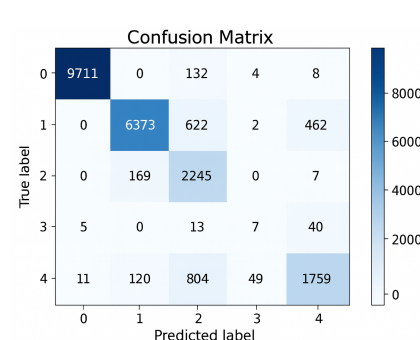


Figure 14. Confusion matrix (RF with normalization)

5.2. Extended AD+DCA process

The integration of prior knowledge in the AD+DCA process provided further benefits. False alerts decreased by up to 90%, the balance of false positives and negatives improved by about 4%, and human intervention was reduced by more than 75%. The framework also handled real-time streams effectively, mitigating drift and data mixture while supporting fraud detection. Finally, an agent-based simulation was conducted with N_d devices (2000–16000), each running 1500 iterations under a similarity threshold $\epsilon = 0.7$. As shown in Figures 15 and 16, precision and recall improved consistently with larger populations, demonstrating coherent clustering, fewer anomalies, and stable accuracy gains over time.

5.3. Comparative summary

Table 4 compares the performance of traditional classifiers, showing that normalization improves their accuracy. The hybrid AD+DCA framework outperforms these methods with a 92% accuracy and a significant reduction in false alarms. This highlights the effectiveness of integrating semantic knowledge to enhance system reliability and robustness.

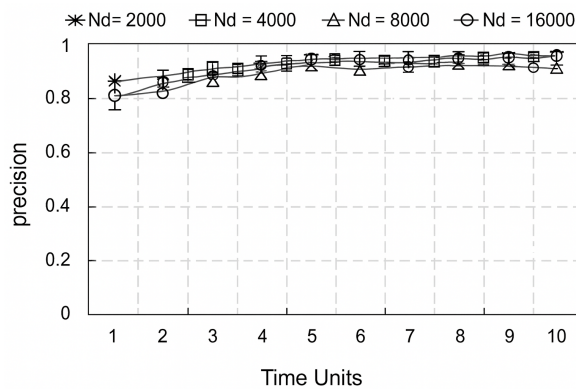


Figure 15. Precision when N_d ranges from 2000 to 16,000

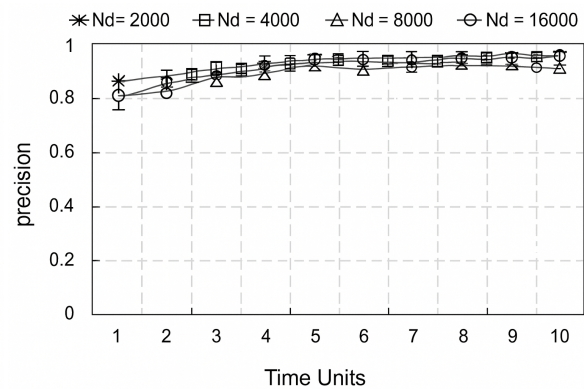


Figure 16. Recall when N_d ranges from 2000 to 16,000

Table 4. Performance comparison of classifiers and AD+DCA framework on NSL-KDD

Model	Accuracy (%)	Precision	Recall	F1-score
GaussianNB (raw)	80	0.79	0.78	0.78
Decision tree (raw)	88	0.87	0.88	0.87
KNN (raw)	73	0.72	0.71	0.71
RF (raw)	90	0.89	0.90	0.89
SVM (raw)	88	0.87	0.88	0.87
GaussianNB (normalized)	75	0.74	0.73	0.73
Decision tree (normalized)	88	0.87	0.88	0.87
KNN (normalized)	87	0.86	0.87	0.86
RF (normalized)	90	0.89	0.90	0.89
SVM (normalized)	88	0.87	0.88	0.87
AD+DCA (proposed)	92	0.91	0.92	0.91

6. DISCUSSION AND ANALYSIS

The results highlight persistent gaps in anomaly detection and RCA. Existing systems often rely only on raw data, which reduces their ability to identify attacks or abnormal behaviors in dynamic contexts. They also suffer from high false alarm rates, rigidity when facing evolving data streams, and frequent manual reconfiguration of RCA models. Added to this are limited transparency in decision-making and weak interoperability across heterogeneous architectures. These factors restrict both scalability and trust in current solutions. By embedding prior knowledge and contextual rules into the detection pipeline, our approach directly addresses these shortcomings. Detection accuracy improved by more than 2% over state-of-the-art

baselines such as RF and SVM, while false positives and negatives dropped by about 4%. Most importantly, spurious alerts were reduced by up to 90%, significantly enhancing the reliability of the system. These improvements are not only statistical but also practical, as they strengthen interpretability and increase confidence in automated results.

Operational benefits are equally notable. Manual interventions for parameter adjustment and threshold tuning were reduced by more than 75%, simplifying deployment in fast-changing environments. The framework also remained stable under streaming conditions, where issues such as temporal drift and mixed data typically hinder existing models. Its applicability extends beyond cybersecurity to industrial monitoring and healthcare, confirming robustness and adaptability. From an AI perspective, the integration of ML with semantic reasoning and bio-inspired multi-agent modeling illustrates a neuro-symbolic approach that balances accuracy with explainability. This hybrid design strengthens anomaly detection as a core AI capability, not just a system integration task. Moreover, the framework is transferable to more recent benchmarks (e.g., CICIDS2017, UNSW-NB15), paving the way for broader validation and generalization.

7. CONCLUSION

This work proposed a hybrid AI framework to improve anomaly detection and data preparation in intelligent systems. The method addresses semantic drift, loss of meaning in large datasets, preprocessing bottlenecks, and data integrity issues. By combining ML with prior knowledge and semantic rules, the framework enables real-time profiling, early correction, and proactive response to anomalies. These advances strengthen both reliability and interpretability while reducing operational risks. Beyond technical contributions, the approach also demonstrates economic impact, with potential cost reductions of up to 50% and performance gains of around 40% over conventional methods. Future extensions will target fully automated big data pipelines, enhanced stress tolerance, regulatory compliance, and advanced repositories to support analysts and decision-makers. Such developments aim to reinforce scalability, resilience, and autonomy in next-generation AI-based security systems.

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Tahri Rachid	✓	✓	✓			✓			✓	✓	✓	✓	✓	✓
Ouammou Abdellah	✓				✓				✓	✓		✓	✓	
Lasbahani Abdellatif		✓	✓		✓					✓	✓			
Abdessamad Jarrar	✓			✓		✓				✓	✓			
Balouki Youssef		✓					✓	✓		✓		✓		

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal Analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project Administration

Fu : Funding Acquisition

CONFLICT OF INTEREST STATEMENT

The authors certify that they have no financial or personal conflicts of interest that could have influenced the work presented in this paper.




DATA AVAILABILITY

The dataset employed in this study (NSL-KDD) is openly accessible from the Canadian Institute for Cybersecurity (CIC), University of New Brunswick: <http://www.unb.ca/cic/datasets/nsl.html>.




REFERENCES

- [1] I. Souiden, Z. Brahmi, and H. Toumi, "A survey on outlier detection in the context of stream mining: Review of existing approaches and recommendations," in *Advances in Intelligent Systems and Computing*, Springer, 2017, pp. 372–383, doi: 10.1007/978-3-319-53480-0_37.
- [2] M. Solé, V. Muntés-Mulero, A. I. Rana, and G. Estrada, "Survey on models and techniques for root-cause analysis," *arXiv:1701.08546*, Jul. 2017.
- [3] W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, and J. Kacprzyk, *Advances in dependability engineering of complex systems: proceedings of the twelfth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX*, Cham, Switzerland: Springer, 2017, doi: 10.1007/978-3-319-59415-6.
- [4] A. Lasbahani and C. Taoussi, "A new unsupervised learning-based process for extraction of knowledge's and improving anomalies detection," in *Journal of Physics: Conference Series*, IOP Publishing, 2021, doi: 10.1088/1742-6596/1743/1/012024.
- [5] H. Huang and S. P. Kasiviswanathan, "Streaming anomaly detection using randomized matrix sketching," *Proceedings of the VLDB Endowment*, vol. 9, no. 3, pp. 192–203, 2016, doi: 10.14778/2850583.2850593.
- [6] J. Jabez and B. Muthukumar, "Intrusion detection system (IDS): Anomaly detection using outlier detection approach," *Procedia Computer Science*, vol. 48, pp. 338–346, 2015, doi: 10.1016/j.procs.2015.04.191.
- [7] S. Ahmad, A. Lavin, S. Purdy, and Z. Agha, "Unsupervised real-time anomaly detection for streaming data," *Neurocomputing*, vol. 262, pp. 134–147, 2017, doi: 10.1016/j.neucom.2017.04.070.
- [8] Y. He and M. Chen, "A probabilistic, mechanism-Independent outlier detection method for online experimentation," in *Proceedings - 2017 International Conference on Data Science and Advanced Analytics, DSAA 2017*, IEEE, 2017, pp. 640–647, doi: 10.1109/DSAA.2017.64.
- [9] T. T. Ademujimi, M. P. Brundage, and V. V. Prabhu, "A review of current machine learning techniques used in manufacturing diagnosis," in *IFIP Advances in Information and Communication Technology*, Springer, 2017, pp. 407–415, doi: 10.1007/978-3-319-66923-6_48.
- [10] B. A. Smith and R. W. Wilkerson, "Fault diagnosis using first order logic tools," in *Midwest Symposium on Circuits and Systems*, IEEE, 1990, pp. 299–302, doi: 10.1109/mwscas.1989.101851.
- [11] H. Paulheim, "Exploiting linked open data as background knowledge in data mining," *CEUR Workshop Proceedings*, vol. 1082, 2013.
- [12] D. Q. Nguyen, "A survey of embedding models of entities and relationships for knowledge graph completion," *Proceedings of the Graph-based Methods for Natural Language Processing (TextGraphs)*, 2019, pp. 1–14.
- [13] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. van den Berg, I. Titov, and M. Welling, "Modeling relational data with graph convolutional networks," *The Semantic Web (ESWC 2018)*, pp. 593–607, 2018, doi: 10.1007/978-3-319-93417-4_38.
- [14] A. M. Raykar and K. B. Ashwini, "A comparative study of machine learning algorithms on intrusion detection system," in *Journal of Machine and Computing*, EDP Sciences, 2022, pp. 67–73, doi: 10.53759/7669/jmc202202009.
- [15] E. Camossi, P. Villa, and L. Mazzola, "Semantic-based anomalous pattern discovery in moving object trajectories," *arXiv:1305.1946*, 2013.
- [16] R. Tahri, Y. Balouki, A. Jarrar, and A. Lasbahani, "Intrusion detection system using machine learning algorithms," *ITM Web of Conferences*, vol. 46, 2022, doi: 10.1051/itmconf/20224602003.
- [17] A. Forestiero, "Metaheuristic algorithm for anomaly detection in Internet of Things leveraging on a neural-driven multiagent system," *Knowledge-Based Systems*, vol. 228, 2021, doi: 10.1016/j.knosys.2021.107241.
- [18] A. Forestiero, "Self-organizing anomaly detection in data streams," *Information Sciences*, vol. 373, pp. 321–336, 2016, doi: 10.1016/j.ins.2016.09.007.
- [19] A. Mniai and K. Jebari, "Credit card fraud detection by improved SVDD," in *Proceedings of the world congress on engineering*, 2022, pp. 6–8.
- [20] G. Song, J. Hu, and H. Wang, "An anti-DoS duplicate address detection model," *Engineering Letters*, vol. 30, no. 2, Jun. 2022.
- [21] A. Ouammou, M. Hanini, S. El Kaffali, and A. Ben Tahar, "Energy consumption and cost analysis for data centers with workload control," in *Innovations in Bio-Inspired Computing and Applications*, Springer, 2018, pp. 92–101, doi: 10.1007/978-3-319-76354-5_9.
- [22] A. Huang, "Similarity measures for text document clustering," in *Proceedings of the Sixth New Zealand Computer Science Research Student Conference (NZCSRSC2008)*, 2008, pp. 49–56.
- [23] R. C. Eberhart, Y. Shi, and J. Kennedy, *Swarm intelligence*. San Francisco, United States: Morgan Kaufmann Publishers, 2001.
- [24] A. Ouammou, M. Hanini, A. B. Tahar, and S. El Kaffali, "Analysis of a M/M/k system with exponential setup times and reserves servers," in *ACM International Conference Proceeding Series*, 2019, pp. 1–5, doi: 10.1145/3372938.3372996.
- [25] H. E. Mohtadi, A. Ouammou, M. Hanini, and A. Haqiq, "Resilient vehicular fog computing networks: an analytical approach to system reliability under breakdown and vacation interruptions," *Cluster Computing*, vol. 28, no. 2, 2025, doi: 10.1007/s10586-024-04805-9.
- [26] C. W. Reynolds, "Flocks, herds, and schools: A distributed behavioral model," in *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1987*, 1987, pp. 25–34, doi: 10.1145/37401.37406.
- [27] A. Ouammou, A. B. Tahar, M. Hanini, and S. E. Kaffali, "Modeling and analysis of quality of service and energy consumption in cloud environment," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 10, no. 9, 2018.
- [28] X. Cui, J. Gao, and T. E. Potok, "A flocking based algorithm for document clustering analysis," *Journal of Systems Architecture*, vol. 52, no. 8–9, pp. 505–515, 2006, doi: 10.1016/j.sysarc.2006.02.003.




BIOGRAPHIES OF AUTHORS

Dr. Tahri Rachid    was born in Zagora, Morocco, in 1990. He obtained a Master's degree in Networks and Information Systems from the Faculty of Sciences and Techniques, Settat, in 2014, and is currently pursuing a Ph.D. at Hassan First University. His research interests focus on the integration of security within artificial intelligence and machine learning systems. He can be contacted at email: rachid.tahrtr@gmail.com.






Dr. Ouammou Abdellah    holds a Ph.D. in Applied Mathematics from the Computer, Networks, Mobility and Modeling Laboratory at Hassan First University of Settat, Morocco. He currently serves as a professor at the Ministry of National Education, Preschool and Sports, while being affiliated with the Faculty of Sciences and Techniques of the same university. His academic interests include probability theory, stochastic and discrete optimization, and their applications in cloud computing. He can be contacted at email: a.ouammou@uhp.ac.ma.






Dr. Lasbahani Abdellatif    is a professor at Sultan Moulay Slimane University and earned his Ph.D. from Hassan First University. His work focuses on formal methods and artificial intelligence, with an emphasis on modeling complex systems. He actively contributes to research through teaching, publications, and collaborations. He can be contacted at email: abbdellatif.lasbahani@gmail.com.



Dr. Abdessamad Jarrar    is a faculty member at Mohammed First University in Oujda, Morocco. He received his Ph.D. from Hassan First University and has more than eight years of research experience. His main interests are formal methods and artificial intelligence, and he regularly publishes and participates in international collaborations. He can be contacted at email: abdessamad.jarrar@gmail.com.



Dr. Balouki Youssef    obtained his Doctor Habilitatus (D.Sc.) in Computer Science from the Faculty of Sciences and Techniques, Settat, Hassan First University. He is a professor and director of graduate studies in computer science, as well as head of the Computational Intelligence research group. His research covers type-2 fuzzy logic, modular neural networks, and neuro-fuzzy systems. He also serves as President of the Academic Association for Computer Science. He can be contacted at email: balouki.youssef@gmail.com.