AIMS Mathematics, 11(7): 23344–23370.
DOI: 10.3934/math.2026941
Received: 25 March 2026
Revised: 07 July 2026
Accepted: 14 July 2026
Published: 31 July 2026

http://www.aimspress.com/journal/Math

Research article

Lie algebra on weighted function spaces for modeling network anomalies

Hamza Alzaareer[1,*] **and Ahmad M. AL-Diabat**[2]

[1] Department of Mathematics, Al-Zaytoonah University of Jordan, Queen Alia Airport St 594, Amman 11733, Jordan

[2] Department of Basic Science, Al-Zaytoonah University of Jordan, Queen Alia Airport St 594, Amman 11733, Jordan

* **Correspondence:** Email: h.alzaareer@zuj.edu.jo.

Abstract: The real-time analysis of dynamic and high-dimensional data streams, such as network traffic, raises a number of mathematical and structural difficulties that are not adequately addressed by purely statistical approaches. This study proposes a theoretical framework for anomaly detection based on the construction of topological Lie algebras on weighted spaces of continuous functions. From a mathematical perspective, network states are modeled as elements of a weighted function space $\mathrm{CV}(X, \mathfrak{g})$, where the underlying set X represents the network domain (for instance, nodes or time indices), the Lie algebra $\mathfrak{g}$ encodes admissible state transitions, and a Nachbin family of weights V reflects the nonuniform relevance of different regions of the network. The core contribution of this paper is the identification of a concrete condition on the weight system V, most notably the multiplicative stability condition $V \preccurlyeq V \cdot V$, which ensures that the pointwise Lie bracket induces a jointly continuous Lie algebra structure on $\mathrm{CV}(X, \mathfrak{g})$. Such an algebraic setting provides a principled mechanism for describing deviations from normal traffic behavior in terms of Lie brackets. In addition, it is shown that the subspace $\mathrm{CV}_0(X, \mathfrak{g})$, corresponding to traffic patterns that vanish at infinity, forms a closed Lie ideal. The closed Lie ideal allows for a natural mathematical distinction between localized anomalies and global, systemwide events. The developed theory is incorporated into an algebraically structured anomaly detection engine (AADE), designed to complement conventional detection techniques by capturing structural aspects of coordinated or evolving attack patterns. The framework provides explicit mathematical characterizations of diverse attack vectors, including distributed denial of service, lateral movement, data exfiltration, and advanced persistent threats, establishing a direct correspondence between algebraic structures and cybersecurity phenomena. As a proof-of-concept validation, experiments on stratified subsets of the NSL-KDD and CIC-IDS-2017 datasets demonstrate that the proposed approach achieves competitive performance compared to both classical machine learning and deep learning baselines. Notably, the algebraic classification mechanism achieves an accuracy of 87.4% in distinguishing between localized and systemic threats, with a false localization rate of approximately 12%. Although the overall detection performance is slightly lower than that of bidirectional long short-term memory (Bi-LSTM) (96.0% vs 96.4% F1-

score), the approach demonstrates notable strength in detecting sophisticated attack patterns such as low-and-slow exfiltration, achieving a detection rate of 68.9% compared to 61.8% for Bi-LSTM. The observed performance suggests that the algebraic structure captures complementary structural information in specific scenarios.

Keywords: Lie algebras; weighted function spaces; cybersecurity; anomaly detection; network traffic analysis; topological algebras; intrusion detection systems
Mathematics Subject Classification: 17B65, 46E25

Nomenclature

Abbreviation	Description
AADE	Algebraically structured anomaly detection engine
CIC-IDS	Canadian institute for cybersecurity - intrusion detection system
CNN	Convolutional neural network
DDoS	Distributed denial of service
DoS	denial of service
FPR	False positive rate
GMM	Gaussian mixture model
LSTM	Long short term memory
NSL-KDD	National security laboratory - knowledge discovery and data mining
R2L	Remote to local
RF	Random forest
SDN	Software defined networking
SVM	Support vector machine
U2R	User to root

1. Introduction: from cyber threat analysis to algebraic formalism

Modern networked systems continuously generate vast streams of diverse data. Extracting actionable security insights from such data, especially identifying malicious behaviors hidden within otherwise legitimate traffic, remains a central challenge in contemporary cybersecurity. The task of anomaly detection has traditionally relied on statistical thresholds or signature-based methods. Although such approaches can be effective against known attack patterns, they often fall short when facing adaptive, coordinated, or slowly evolving threats designed to mimic normal traffic statistics. Machine learning approaches for intrusion detection have been extensively surveyed by Al Lail et al. [1], and empirical comparisons of botnet detection methods were provided by Garcia et al. [2].

A different perspective arises when network traffic is considered not merely as a collection of numerical measurements but as a dynamic system with inherent structure. In typical operational environments, normal traffic displays recurring interaction patterns, temporal regularities, and stable transitions between states. Malicious activities often manifest as subtle distortions of such patterns

rather than abrupt deviations in numerical metrics. Capturing structural changes therefore requires a mathematical framework capable of representing transformations and their interactions.

Lie algebras offer exactly such a framework. Originally developed to formalize infinitesimal symmetries in geometry and physics, Lie algebras have become a central tool in the study of dynamical systems and continuous transformations. The foundational theory of Lie groups and algebras was established by Bourbaki [3], with further developments in symmetries and representations by Fuchs and Schweigert [4] and by de Azcárraga and Izquierdo [5]. In the current work, network traffic evolution is proposed to be modeled using functions that take values in a Lie algebra. Such a shift enables anomaly detection to focus not only on statistical deviations but also on structural and algebraic changes in the traffic.

More specifically, the network domain, which may include nodes, IP addresses, or time intervals, is represented as a topological space X. The instantaneous state of network traffic, abstracted through features such as flow rates, protocol types, or connection patterns, is modeled as an element of a locally convex space. By framing traffic in such a way, a foundation is provided for analyzing its evolution using algebraic and topological tools, enabling more nuanced analysis of anomalies that may be difficult to characterize through conventional statistical methods alone.

Not all regions of the network, however, are of equal relevance from a security perspective. Certain nodes or time windows may warrant closer scrutiny due to their strategic importance or historical vulnerability. Such asymmetry is encoded by a family of weight functions V on X, where larger values indicate higher monitoring priority.

Within the current framework, the object of study becomes the space $\mathrm{CV}(X, \mathfrak{g})$ of continuous functions from X into a Lie algebra $\mathfrak{g}$, with continuity controlled by the weights in V. Weighted spaces of such type were introduced by Nachbin and have been systematically studied by Bierstedt, Prolla, Summers, and others in the context of functional analysis. Bierstedt [6] investigated weighted spaces of vector-valued functions, Prolla [7] explored their approximation properties, and Summers [8] addressed complex bounded cases. The choice of the Lie algebra $\mathfrak{g}$ reflects the admissible infinitesimal transitions between traffic states; for example, matrix Lie algebras such as $\mathfrak{gl}(n, \mathbb{R})$ can be used to model interactions among a finite collection of traffic profiles.

The central mathematical question addressed in the paper is the following: Under which condition does the weighted space $\mathrm{CV}(X, \mathfrak{g})$ itself carry a natural structure of a topological Lie algebra, with a jointly continuous pointwise bracket? Such a question is not merely of abstract interest. Continuity of the Lie bracket is essential for ensuring that small perturbations in observed traffic lead to controlled variations in the resulting anomaly indicators, a property that is crucial for numerical stability in real-time detection systems.

Recent developments in algebraic and structural anomaly detection further motivate the current approach. For instance, Lie algebra methods have been successfully applied to data analysis and machine learning by Said et al. [9], and algebraic frameworks have been proposed for intrusion detection in software-defined networks by Kang et al. [10]. Similarly, structural anomaly detection in time series using Lie groups was studied by David and Gu [11]. Additionally, interaction-aware methods via algebraic graph theory were developed by Akoglu et al. [12], highlighting the growing relevance of algebraic structures in cybersecurity. Building on such advances and the foundational theory of Lie algebras, the current work introduces a topological Lie algebra framework on weighted function spaces specifically tailored for network anomaly detection, offering a foundational and

mathematically rigorous model for capturing both local and global traffic deviations.

Furthermore, advanced control-theoretic and algebraic approaches have been successfully applied to complex networked systems, demonstrating the critical role of structural modeling in managing complex network behaviors. For instance, event-triggered input strategies for state-constrained containment control in nonlinear multiagent systems were studied by Li et al. [13], and optimized adaptive finite-time consensus control for stochastic systems with non-affine faults was developed by Zhang et al. [14], highlighting the effectiveness of nonstatistical, dynamic frameworks. Inspired by such structural approaches, the current work further solidifies the paradigm of moving beyond purely statistical methods to capture the deep algebraic interactions within network traffic.

The main result demonstrates that the desired property holds precisely when the weight family V satisfies a multiplicative stability condition, expressed as $V \preccurlyeq V \cdot V$. Such a requirement emerges naturally from the need to control interactions between weighted traffic models under the Lie bracket operation. With such a structure in place, it becomes possible to mathematically distinguish between localized, transient events and persistent, systemwide anomalies. In particular, the subspace $\mathrm{CV}_0(X, \mathfrak{g})$, consisting of functions that vanish at infinity, forms a closed Lie ideal. The framework provides a principled and mathematically consistent foundation for classifying different types of network threats.

The current work establishes the theoretical foundations and demonstrates feasibility through proof-of-concept experiments on benchmark datasets. Comprehensive empirical validation across diverse network configurations, alternative algebraic structures, and large-scale deployments is reserved for future work.

The structure of the paper is as follows. Section 2 presents the necessary background on weighted function spaces and reviews relevant concepts from Lie theory. In Section 3, the main theoretical results are developed, establishing the Lie algebra structure on $\mathrm{CV}(X, \mathfrak{g})$ and proving the ideal property of $\mathrm{CV}_0(X, \mathfrak{g})$. Section 4 translates such abstract results into a practical, application-oriented setting by outlining an anomaly detection architecture grounded in algebraic principles and discussing its advantages compared to classical approaches. Section 5 presents the experimental validation of the proposed framework. Section 6 provides a sensitivity analysis and discusses the impact of parameter settings on the detection performance. Finally, Section 7 summarizes the main contributions and outlines directions for future research, including the extension to spaces of higher regularity, the investigation of inductive limits, and algorithmic implementation.

2. Mathematical preliminaries and model specification

Weighted spaces of continuous functions were first introduced by Nachbin [15] and further studied by Bierstedt [6], Prolla [7], Summers [8], and Oubbi [16], who later extended such results to vector-valued functions [17]. These spaces offer a versatile analytical framework for incorporating spatial or temporal priorities via weight functions.

The current work extends these ideas to weighted spaces of Lie algebra-valued functions, leveraging their algebraic structure to provide a rigorous foundation for anomaly detection in network traffic. The proposed approach allows for the capture of both the structural and algebraic aspects of traffic evolution, which are often overlooked by conventional statistical methods.

The current section establishes the mathematical objects at the heart of the model, tailoring standard definitions from weighted analysis and Lie theory to the cybersecurity context.

2.1. Network domain and feature space

Let X be a completely regular Hausdorff space. In practice, X can be any of the following:

- A finite or countably infinite set of network nodes or IP addresses, endowed with the discrete topology.
- A compact time interval $[0, T]$ for analyzing a specific session, endowed with the standard Euclidean topology.
- The noncompact time axis $\mathbb{R}_{\geq 0}$ for continuous monitoring, endowed with the standard Euclidean topology.
- A product space Nodes $\times$ Time for spatiotemporal analysis, endowed with the product topology.

The topology on X allows for the definition of continuous functions, which model traffic patterns that evolve continuously over the network domain.

Let E be a locally convex topological vector space. In applications, E is typically $\mathbb{R}^n$ or $\mathbb{C}^n$, where each coordinate represents a normalized feature of network traffic at a point $x \in X$ (e.g., bytes per second, packets per second, number of active connections, flag counts). The topology on E is given by a family $\mathbb{P}$ of seminorms that controls both.

Monitoring resources are finite. A security policy must prioritize some network segments over others. Such prioritization is encoded mathematically via weights.

Definition 1 (Weights and Nachbin family of weights)**.** *A weight on X is an upper semicontinuous function $v\colon X \to [0, \infty)$. A Nachbin family or system of weights V is a set of weights satisfying the following:*
(1) *For every $x \in X$, there exists $v \in V$ with $v(x) > 0$ (pointwise positivity).*
(2) *For any $v_1, v_2 \in V$, there exists $v \in V$ such that $v_1 \leq v$, and $v_2 \leq v$ (directedness), where $\leq$ denotes the pointwise partial order of functions.*

Nachbin families allow for the modeling of priority and localization effects and were introduced in [15].

Interpretation: A weight v assigns a non-negative priority coefficient to each location $x \in X$. Upper semicontinuity means that if $v(x_0)$ is high, it remains high in a neighborhood of x_0, which aligns with the idea that critical assets are often grouped. Condition (1) that ensures every location is monitored by some weight. Condition (2) allows for the combination of security policies; if one policy gives weight v_1 and another v_2, there is a unified policy v that respects both. For example, if $V = \{v\}$ consists of a single weight, it forms a valid Nachbin family as long as $v(x) > 0$ for all $x \in X$.

2.2. Weighted spaces of traffic models

The primary spaces for modeling traffic are now defined. Let $\mathrm{C}(X, E)$ denote the space of all continuous functions from X to E.

Definition 2 (Weighted traffic model spaces)**.** *For a given Nachbin family V, the following spaces are defined:*

$$\mathrm{CV}(X, E) := \Big\{ f \in \mathrm{C}(X, E) \;\Big|\; \forall v \in V,\ x \mapsto v(x)\, p(f(x)) \text{ is bounded on } X \text{ for all } p \in \mathbb{P} \Big\}, \tag{2.1}$$

$$\mathrm{CV}_0(X, E) := \Big\{ f \in \mathrm{C}(X, E) \;\Big|\; \forall v \in V,\ x \mapsto v(x)\, p(f(x)) \text{ vanishes at infinity on } X,\ \forall p \in \mathbb{P} \Big\}. \tag{2.2}$$

These spaces were studied extensively in [6, 7, 17].

Interpretation:

- CV(X, E): Represents the space of all globally defined traffic models. The condition

$$x \mapsto v(x)\, p(f(x))$$

bounded means that even high-intensity traffic (large $p(f(x))$) is acceptable in the model if it occurs in a low-priority area (small $v(x)$). The bounded condition reflects resource-aware modeling.

- $\mathrm{CV}_0(X, E)$: Represents the space of localized traffic events. "Vanishes at infinity" means that outside of some compact (i.e., finite) subset of X, the weighted traffic intensity becomes arbitrarily small. The vanishing property models short-duration attacks, flash crowds, or any activity confined to a limited part of the network.

These spaces are equipped with the weighted topology, generated by the seminorms:

$$\mathrm{P}_{v,p}(f) = \sup_{x\in X} v(x)\, p(f(x)), \qquad v \in V,\ p \in \mathbb{P}. \tag{2.3}$$

The weighted topology makes CV(X, E) and $\mathrm{CV}_0(X, E)$ Hausdorff locally convex spaces. Importantly, $\mathrm{CV}_0(X, E)$ is a closed subspace of CV(X, E).

Let $(\mathfrak{g}, [\cdot, \cdot]_{\mathfrak{g}})$ be a locally convex topological Lie algebra. This means that $\mathfrak{g}$ is a locally convex space, and the Lie bracket $[\cdot, \cdot]_{\mathfrak{g}} \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a jointly continuous bilinear map satisfying skew-symmetry and the Jacobi identity.

The choice of $\mathfrak{g}$ is application-dependent and crucial. Two illustrative examples are:

(1) Matrix algebra $\mathfrak{gl}(n, \mathbb{R})$: Here, $\mathfrak{g}$ consists of all $n \times n$ real matrices with the commutator bracket $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. This can model a system with n discrete "traffic states" (e.g., idle, normal transfer, suspicious scan, flood). A matrix $\mathbf{A} \in \mathfrak{gl}(n, \mathbb{R})$ acts as an infinitesimal generator of transition rates between these states.

(2) Algebra of vector fields Vect(M): If the space of normal traffic profiles is modeled as a smooth manifold M, then $\mathfrak{g} = \mathrm{Vect}(M)$, the Lie algebra of smooth vector fields on M. An element $\gamma \in \mathrm{CV}(X, \mathrm{Vect}(M))$ assigns to each network point x a direction in which the traffic profile on M is instantaneously changing.

For the core theory, $\mathfrak{g}$ is kept general, requiring only that it is a topological Lie algebra.

The central object is the weighted space of Lie algebra-valued functions, $\mathrm{CV}(X, \mathfrak{g})$. An element $\gamma \in \mathrm{CV}(X, \mathfrak{g})$ is a continuous map that assigns to each network location x an infinitesimal generator of change $\gamma(x) \in \mathfrak{g}$. One can think of γ as a field of transformation generators over the network.

On $\mathrm{CV}(X, \mathfrak{g})$, a pointwise Lie bracket is defined:

$$[\gamma, \eta](x) := [\gamma(x), \eta(x)]_{\mathfrak{g}}, \qquad \forall x \in X,\ \forall \gamma, \eta \in \mathrm{CV}(X, \mathfrak{g}). \tag{2.4}$$

The pointwise bracket is the natural candidate for a Lie algebra structure on the function space. The pivotal question is: When is this bracket a continuous operation with respect to the weighted topology? The answer, provided in the next section, hinges on a compatibility condition between the weight family V and the multiplication of functions.

3. Main theoretical results

This section presents the core theorems that transform the weighted function space into a structured environment suitable for algorithmic analysis. The proofs, although mathematically rigorous, are motivated by the need for stability in computation.

Definition 3 (Multiplicative stability)**.** *Let V be a Nachbin family on X. The product*

$$V \cdot V := \{v \cdot w \mid v, w \in V\}$$

is defined, where $(v \cdot w)(x) = v(x)w(x)$. The family V is said to be multiplicatively stable (or to satisfy the product condition) if $V \leq V \cdot V$; that is for every $v \in V$, there exist $w_1, w_2 \in V$ such that $v \leq w_1 \cdot w_2$. More generally, this can be relaxed to $V \preccurlyeq V \cdot V$, meaning that for every $v \in V$, there exist $w_1, w_2 \in V$ and $\mu > 0$ such that $v \leq \mu w_1 \cdot w_2$.

Remark 4. *Note that $V \cdot V$ itself forms a Nachbin family. Indeed, for any $x \in X$, there exists $v \in V$ with $v(x) > 0$, so $v^2 \in V \cdot V$ satisfies $v^2(x) > 0$. For directedness, if $u_i = v_i \cdot w_i \in V \cdot V$ for $i = 1, 2$, by the directedness of V, there exist $v, w \in V$ such that $v_i \leq v$, and $w_i \leq w$; hence, $u_i \leq v \cdot w \in V \cdot V$. Furthermore, in the special case where $V = \{v\}$ consists of a single weight, $V \cdot V = \{v^2\}$. In this case, the condition $V \leq V \cdot V$ holds if and only if $v(x) \geq 1$ for all $x \in X$, whereas the relaxed condition $V \preccurlyeq V \cdot V$ holds if and only if* $\inf_{x \in X} v(x) > 0$.

The multiplicative stability condition is the linchpin of the theory. It ensures that the "interaction" of two traffic models, as measured by their Lie bracket, does not produce a function that grows too fast relative to the weights. In practical terms, it means the weighting scheme must be designed such that the priority assigned to an interaction is at most as high as the product of the priorities assigned to the individual factors at that point. The multiplicative stability is a natural requirement for maintaining control over bilinear operations.

Theorem 5 (Topological Lie algebra structure)**.** *Let X be a completely regular Hausdorff space, V a Nachbin family of weights on X, and $(\mathfrak{g}, [\cdot, \cdot]_{\mathfrak{g}})$ a locally convex topological Lie algebra. If V is multiplicatively stable ($V \leq V \cdot V$), then the weighted space* $\mathrm{CV}(X, \mathfrak{g})$*, equipped with the pointwise Lie bracket and the weighted topology, is a topological Lie algebra. That is, the bracket operation*

$$[\cdot, \cdot] \colon \mathrm{CV}(X, \mathfrak{g}) \times \mathrm{CV}(X, \mathfrak{g}) \to \mathrm{CV}(X, \mathfrak{g}) \tag{3.1}$$

is jointly continuous.

Proof. Let $\gamma, \eta \in \mathrm{CV}(X, \mathfrak{g})$. It must be shown that the bracket is well-defined (i.e., $[\gamma, \eta] \in \mathrm{CV}(X, \mathfrak{g})$) and continuous.

Well-definedness and estimate: Fix a weight $v \in V$ and a continuous seminorm P on $\mathfrak{g}$. By the multiplicative stability of V, there exist weights $w_1, w_2 \in V$ such that $v \leq w_1 \cdot w_2$. Because the bracket in $\mathfrak{g}$ is continuous, there exists a continuous seminorm Q on $\mathfrak{g}$ and a constant $C > 0$ such that for all $a, b \in \mathfrak{g}$, $\mathrm{P}([a, b]_{\mathfrak{g}}) \leq C\ \mathrm{Q}(a)\ \mathrm{Q}(b)$. (For many Lie algebras, like matrix algebras, one can take $C = 1$ and Q as the operator norm).

Now, for any $x \in X$, the following holds:

$$\begin{aligned} v(x)P([\gamma,\eta](x)) &\le w_1(x)\, w_2(x)\cdot C\, Q(\gamma(x))\, Q(\eta(x)) \\ &= C\,(w_1(x)Q(\gamma(x)))(w_2(x)Q(\eta(x))) \\ &\le C\,(\sup_{y\in X} w_1(y)Q(\gamma(y)))(\sup_{z\in X} w_2(z)Q(\eta(z))) \\ &= C\, Q_{w_1}(\gamma)\, Q_{w_2}(\eta), \end{aligned} \tag{3.2}$$

where

$$\mathrm{Q}_w(f) := \sup_{x\in X} w(x)\ \mathrm{Q}(f(x))$$

is a defining seminorm for the weighted topology on $\mathrm{CV}(X,\mathfrak{g})$.

Taking the supremum over $x \in X$ on the left-hand side yields the fundamental estimate

$$\mathrm{P}_v([\gamma,\eta]) := \sup_{x\in X} v(x)\ \mathrm{P}([\gamma,\eta](x)) \leqslant C\ \mathrm{Q}_{w_1}(\gamma)\ \mathrm{Q}_{w_2}(\eta) < \infty. \tag{3.3}$$

The finiteness follows because $\gamma, \eta \in \mathrm{CV}(X,\mathfrak{g})$, so $Q_{w_1}(\gamma)$ and $Q_{w_2}(\eta)$ are finite. Because v and P are arbitrary, this proves $[\gamma,\eta] \in \mathrm{CV}(X,\mathfrak{g})$.

Continuity: Estimate (3.3) shows that the bracket is continuous at $(0,0)$. For bilinear maps on locally convex spaces, continuity at the origin implies joint continuity everywhere. Therefore, the Lie bracket is jointly continuous on $\mathrm{CV}(X,\mathfrak{g})$. □

Theorem 5 provides the mathematical foundation. The space of traffic models is not just a set; it is an algebra where the "multiplication" (the Lie bracket) is a stable, continuous operation. Such stability is crucial for any algorithm that computes with these brackets, as small errors in input models will lead to controlled errors in the output.

Theorem 6 (Localized events form an ideal)**.** *Under the same hypotheses as Theorem 5, in particular having $V \le V \cdot V$, the subspace $\mathrm{CV}_0(X,\mathfrak{g})$ is a closed Lie ideal of $\mathrm{CV}(X,\mathfrak{g})$. That is:*
(1) *$\mathrm{CV}_0(X,\mathfrak{g})$ is a Lie subalgebra.*
(2) *For any $\gamma \in \mathrm{CV}(X,\mathfrak{g})$ and any $\eta \in \mathrm{CV}_0(X,\mathfrak{g})$, the bracket $[\gamma,\eta] \in \mathrm{CV}_0(X,\mathfrak{g})$.*

Proof. Standard functional analysis establishes that $\mathrm{CV}_0(X,\mathfrak{g})$ is a closed subspace. The proof that it is an ideal follows. Let $\gamma \in \mathrm{CV}(X,\mathfrak{g})$ and $\eta \in \mathrm{CV}_0(X,\mathfrak{g})$. Fix $v \in V$ and a seminorm P on $\mathfrak{g}$. As before, find $w_1, w_2 \in V$ with $v \le w_1 \cdot w_2$ and a seminorm Q on $\mathfrak{g}$ with

$$v(x)\ \mathrm{P}([\gamma,\eta])(x) \le C\, w_1(x)\ \mathrm{Q}(\gamma(x))\, w_2(x)\ \mathrm{Q}(\eta(x)).$$

Because $\eta \in \mathrm{CV}_0(X,\mathfrak{g})$, for the weight w_2 and seminorm Q and for any $\epsilon > 0$, there exists a compact set $K \subset X$ such that

$$w_2(x)\ \mathrm{Q}(\eta(x)) < \frac{\varepsilon}{C\,\mathrm{Q}_{w_1}(\gamma) + 1} \quad \text{for all } x \notin K. \tag{3.4}$$

(If $Q_{w_1}(\gamma) = 0$, the denominator is simply 1.)

Now, for $x \notin K$, the estimate from the previous proof applies:

$$\begin{aligned} v(x)\,\mathrm{P}([\gamma,\eta](x)) &\le C\,(w_1(x)\,\mathrm{Q}(\gamma(x)))(w_2(x)\,\mathrm{Q}(\eta(x))) \\ &< C\,\mathrm{Q}_{w_1}(\gamma)\cdot\frac{\varepsilon}{C\ \mathrm{Q}_{w_1}(\gamma)+1} \\ &< \varepsilon. \end{aligned} \tag{3.5}$$

The preceding inequality proves that the function $x \mapsto v(x)\ \mathrm{P}([\gamma, \eta](x))$ is less than ϵ outside the compact set K. Because $\epsilon > 0$ was arbitrary, this function vanishes at infinity. As v and P are arbitrary, the conclusion $[\gamma, \eta] \in \mathrm{CV}_0(X, \mathfrak{g})$ follows. □

Interpretation for cybersecurity: Theorem 6 provides a clean mathematical classification. The full algebra $\mathrm{CV}(X, \mathfrak{g})$ contains models for any networkwide traffic pattern. The ideal $\mathrm{CV}_0(X, \mathfrak{g})$ contains models for activity that is localized in space or time (vanishes at infinity). The theorem guarantees that the interaction (bracket) of a global pattern with a localized event always yields another localized event. Consequently, localized anomalies cannot, through their interaction with the global state alone, instantly produce a global threat. The ideal property has important triage implications: An alert triggered by an element in CV_0 can be initially treated as a contained incident.

4. Application: algebraically structured anomaly detection

The theoretical framework of Sections 2 and 3 was developed with a clear application in mind. The framework is now synthesized into a proposed system architecture: the algebraically structured detection engine (AADE).

The AADE operates in three sequential phases.

Phase 1: Preprocessing and feature extraction. Network packet headers and flows are processed in real time over sliding windows. For each window and for each logical unit in the domain X (e.g., for each server subnet), a feature vector in $E \simeq \mathbb{R}^n$ is extracted. This vector is normalized with respect to the standard Euclidean norm on $\mathbb{R}^n$.

Phase 2: Embedding into the Lie algebra. This is the crucial, model-dependent step. A map

$$\varphi \colon E \to \mathfrak{g}$$

is needed to translate a feature vector into an infinitesimal generator. For the matrix algebra example $\mathfrak{g} = \mathfrak{gl}(m, \mathbb{R})$, one possible method is as follows:

(1) Cluster historical normal traffic into m representative profiles (states) $S_1, \ldots, S_m$.

(2) For a new feature vector $\mathbf{f} \in E$, compute its affinity (e.g., smoothed inverse distance) to each state S_i, producing a vector $\mathbf{a} \in \mathbb{R}^m$.

(3) Construct a transition rate matrix $\mathbf{A} \in \mathfrak{gl}(m, \mathbb{R})$ where the off-diagonal entry A_{ij} represents the inferred instantaneous rate of transition from state S_i to state S_j given $\mathbf{f}$. The diagonal is set so that rows sum to zero. The resulting matrix $\mathbf{A}$ is the element $\varphi(\mathbf{f}) \in \mathfrak{g}$.

Thus, over a window of time and across the network, a function $\gamma_{\mathrm{obs}} \colon X \to \mathfrak{g}$ is built. By ensuring that the construction respects the weighted boundedness conditions (e.g., by ensuring the feature extraction inherently bounds the outputs, which are then scaled by the specific active weight $v \in V$ corresponding to the node's current security policy), the result is $\gamma_{\mathrm{obs}} \in \mathrm{CV}(X, \mathfrak{g})$.

Phase 3: Weight assignment. The Nachbin family V is not static; it encapsulates a menu of possible security policies and evolves over time. In practice, the system utilizes a sequence of weight functions v_{t_i} indexed by time intervals, reflecting dynamic threat intelligence and asset criticality. Considering a whole family V rather than a single fixed weight allows the framework to adapt to different temporal security policies and spatial priorities simultaneously. For example, if a new vulnerability is announced, the active weight function for nodes running that service is increased, focusing the algebraic analysis on the most critical areas at that specific time.

Operationally, at each time step t, the system selects the active weight $v_{t_i} \in V$ based on current security policy and computes the weighted anomaly score $s(x,t) = v_{t_i}(x) \cdot \|\Delta(x,t)\|_F$, enabling dynamic prioritization of monitoring resources across both spatial and temporal dimensions. Here, $\|\cdot\|_F$ denotes the standard Frobenius norm on the matrix Lie algebra $\mathfrak{g}$, defined for any $\mathbf{A} \in \mathfrak{gl}(n, \mathbb{R})$ as

$$\|\mathbf{A}\|_F = \sqrt{\mathrm{tr}(A^T A)} = \sqrt{\sum_{i,j} |a_{ij}|^2}.$$

This norm is particularly suitable for our framework because it is unitarily invariant and provides a natural measure of the magnitude of algebraic distortions in the transition rate matrices.

The core detection logic of the AADE leverages the Lie algebra structure. A baseline model of "normal" traffic, $\gamma_{\mathrm{norm}} \in \mathrm{CV}(X, \mathfrak{g})$, is maintained. The baseline could be a periodic average, a slowly evolving model learned online, or a set of profiles for different times of day.

The anomaly signal Δ is not merely a difference in norms but is defined using the Lie bracket

$$\Delta := [\gamma_{\mathrm{obs}}, \gamma_{\mathrm{norm}}] \in \mathrm{CV}(X, \mathfrak{g}). \tag{4.1}$$

Why the bracket? The bracket $[\mathbf{A}, \mathbf{B}]$ measures the extent to which the transformations $\mathbf{A}$ and $\mathbf{B}$ fail to commute. In the current context, γ_{obs} represents the instantaneous transformation suggested by the observed traffic. γ_{norm} represents the transformation expected during normal operation. If the observed traffic is merely a scaled or slightly noisy version of the normal one, then γ_{obs} and γ_{norm} will be nearly proportional, and their bracket will be small. However, if the observed traffic is executing a fundamentally different type of transformation—for example, a port scan (which sequentially probes states) versus normal client-server exchanges (which oscillate between a few states)—their generators will not commute, resulting in a large bracket.

The weighted seminorm $\mathrm{P}_{v,p}(\Delta)$ then provides a scalar anomaly score that is already prioritized by the weight v. Theorem 5 guarantees that this score is stable: Small perturbations in γ_{obs} or γ_{norm} lead to small changes in Δ.

Once an anomaly Δ is detected, the AADE classifies it using Theorem 6. The engine checks whether Δ lies in the ideal $\mathrm{CV}_0(X, \mathfrak{g})$. The classification is done operationally by verifying if the function $x \mapsto v(x)\ \mathrm{p}(\Delta(x))$ decays rapidly outside some identifiable compact region K (e.g., a single subnet, a short time burst).

- **Case 1:** $\Delta \in \mathrm{CV}_0(X, \mathfrak{g})$**.** The anomaly is classified as a localized event. Examples: a distributed denial of service (DDoS) attempt against a single server that fails to propagate, a misconfigured client generating unusual but contained traffic. The response may be automated containment on the affected segment.
- **Case 2:** $\Delta \notin \mathrm{CV}_0(X, \mathfrak{g})$**.** The anomaly is classified as a persistent or propagating threat. Examples: a worm spreading laterally across subnets, a coordinated low-and-slow exfiltration from multiple nodes. The systemic nature of the threat triggers a higher-level incident response, as the threat is systemic.

4.1. Mathematical characterization of cybersecurity phenomena

To bridge the gap between the abstract topological framework and real-world network traffic behavior, it is essential to establish a rigorous correspondence between specific cybersecurity

phenomena and their representations within the weighted Lie algebra $\mathrm{CV}(X, \mathfrak{g})$. Although the experimental validation utilizes $\mathfrak{g} = \mathfrak{gl}(4, \mathbb{R})$, a simplified 3-dimensional subsystem is considered in the following illustrative examples to maintain notational clarity. Let the states be S_1 (Idle/Background), S_2 (Active/Normal), and S_3 (Overloaded/Anomalous).

4.1.1. DDoS attacks

A DDoS attack is characterized by a massive influx of traffic directed at a specific target, forcing the system state into an overloaded condition. Mathematically, at the target node $x_{\text{target}} \in X$, the observed generator $\gamma_{\text{obs}}(x_{\text{target}})$ exhibits large transition rates towards S_3, deviating fundamentally from the normal oscillatory behavior between S_1 and S_2.

For illustration, the normal and observed generators at x_{target} can be represented as

$$\gamma_{\text{norm}}(x_{\text{target}}) = \begin{pmatrix} -0.3 & 0.2 & 0.1 \\ 0.1 & -0.3 & 0.2 \\ 0.2 & 0.1 & -0.3 \end{pmatrix}, \qquad \gamma_{\text{obs}}(x_{\text{target}}) = \begin{pmatrix} -0.1 & 0.1 & 0.0 \\ 0.0 & -0.9 & 0.9 \\ 0.8 & 0.0 & -0.8 \end{pmatrix}. \tag{4.2}$$

Both matrices are valid elements of $\mathfrak{gl}(3, \mathbb{R})$ (row sums equal zero). The Lie bracket

$$\Delta(x_{\text{target}}) = [\gamma_{\text{obs}}, \gamma_{\text{norm}}]$$

yields a matrix with large nonzero entries, reflecting the noncommutativity of the forced state transition versus natural traffic oscillation.

Crucially, this algebraic distortion is highly localized. The anomaly signal $\Delta(x)$ is significant only in a small compact neighborhood K around x_{target} and vanishes rapidly for $x \notin K$. Consequently, the Concentration Index $C(K, v)$ (see Definition 7 in Section 5.5) yields $C(K, v) \gg \theta = 0.85$, correctly classifying the DDoS attack as a localized event ($\Delta \in \mathrm{CV}_0(X, \mathfrak{g})$).

4.1.2. Lateral movement

Lateral movement involves an adversary progressively compromising multiple nodes within a network. Unlike a DDoS attack, the algebraic distortion does not vanish at infinity; instead, it propagates across the spatial domain X.

If the attacker moves through a sequence of nodes $x_1, x_2, \ldots, x_k$, the observed generator $\gamma_{\text{obs}}(x_i)$ at each compromised node will exhibit anomalous transition rates (e.g., unauthorized port scanning or credential harvesting transitions). Consequently, the Lie bracket $\Delta(x_i)$ remains significantly nonzero across this distributed set of nodes. Because the support of Δ spans a large, noncompact subset of the network, the weighted traffic intensity does not vanish outside any small compact set K. This results in a low concentration index ($C(K, v) < \theta$), mathematically classifying lateral movement as a systemic threat ($\Delta \notin \mathrm{CV}_0(X, \mathfrak{g})$).

4.1.3. Data exfiltration (low-and-slow)

Data exfiltration attacks are designed to mimic normal background traffic by maintaining a low, steady outflow of data. Statistically, the mean and variance of the feature vector remain largely unchanged, rendering traditional threshold-based detection ineffective. However, the Lie algebraic framework captures the subtle structural distortion in the transition dynamics.

Let the normal generator be γ_{norm}. An exfiltration attack introduces a subtle perturbation ϵE, where E represents the exfiltration transition pattern, and $\epsilon \ll 1$ is a small scaling factor to avoid detection. The observed generator is $\gamma_{\text{obs}} = \gamma_{\text{norm}} + \epsilon E$. The anomaly signal is given by the Lie bracket

$$\Delta = [\gamma_{\text{norm}} + \epsilon E, \gamma_{\text{norm}}] = [\gamma_{\text{norm}}, \gamma_{\text{norm}}] + \epsilon[E, \gamma_{\text{norm}}] = \epsilon[E, \gamma_{\text{norm}}]. \tag{4.3}$$

Although the Euclidean distance $\|\gamma_{\text{obs}} - \gamma_{\text{norm}}\| = \epsilon\|E\|$ is negligibly small (evading statistical detection), the Lie bracket $\Delta = \epsilon[E, \gamma_{\text{norm}}]$ is structurally nonzero as long as the exfiltration pattern E does not commute with the normal traffic pattern γ_{norm}. This demonstrates mathematically why the noncommutative nature of the Lie bracket provides superior sensitivity to low-and-slow structural anomalies compared to purely statistical distance metrics.

4.1.4. Advanced persistent threats (APTs)

APTs represent complex, multistage attacks that evolve over extended periods. In the proposed framework, an APT is modeled as a time-dependent trajectory within the Lie algebra, $\gamma_{\text{obs}}(x, t)$.

The anomaly is not captured by a single instantaneous bracket but by the evolution of the algebraic structure over time. Note that, unlike the static anomaly signal $\Delta(x) = [\gamma_{\text{obs}}(x), \gamma_{\text{norm}}(x)]$ defined in Eq (4.1) for instantaneous detection, the APT's context requires a temporal extension $\Delta(x, t) = [\gamma_{\text{obs}}(x, t), \gamma_{\text{norm}}(x, t)]$, where the time parameter t captures the slow, continuous drift of the attacker's trajectory within the Lie algebra. The Frobenius norm $\|\cdot\|_F$ is used to quantify the magnitude of this time-evolving algebraic distortion. The temporal derivative of the anomaly signal, $\partial/\partial t\Delta(x, t)$, and higher-order Lie brackets (e.g., $[\Delta, [\gamma_{\text{obs}}, \gamma_{\text{norm}}]]$) become critical indicators.

An APT typically exhibits a slow, continuous drift in the Lie algebra, where the trajectory $\gamma_{\text{obs}}(x, t)$ gradually diverges from the normal manifold. The persistent, evolving distortion ensures that $\Delta(x, t)$ does not vanish over time, firmly classifying APTs as persistent systemic threats that require continuous algebraic monitoring rather than static thresholding.

The AADE framework, grounded in the Lie algebraic formulation, provides several unique benefits that distinguish it from conventional anomaly detection methods.

(1) Structural insight beyond statistics: Traditional methods often rely heavily on basic statistical summaries, which may overlook subtle or coordinated anomalies. An AADE, by analyzing the commutative relationships within network traffic transformations, can uncover complex patterns that leave standard metrics such as mean and variance largely unchanged but alter the interactions between events in meaningful ways.

(2) Intelligent resource allocation: The dynamic weight set V adapts to both policy constraints and real-time threat intelligence, directing computational resources to where they are most needed. This ensures efficient monitoring without overwhelming the system with unnecessary processing.

(3) Mathematical consistency and predictable behavior: Based on a rigorous topological algebraic framework (see Theorems 5 and 6), the AADE system guarantees well-defined operations with strict continuity. Consequently, its algorithms behave consistently and predictably under varying network conditions.

(4) Clear and natural threat classification: The ideal structure inherently separates anomalies into localized versus systemic types. This classification provides actionable guidance for responding to threats while remaining grounded in a mathematically justified framework.

5. Experimental validation: proof-of-concept implementation

Although the theoretical framework establishes the mathematical foundations for Lie algebra-based anomaly detection, its practical utility must be validated empirically. This section presents a proof-of-concept implementation on the NSL-KDD dataset [18], a standard benchmark for network intrusion detection, demonstrating feasibility, computational tractability, and comparative advantages over conventional statistical approaches. The authors emphasize that this implementation serves as a *conceptual bridge* between abstract theory and practical deployment rather than a production-ready system.

5.1. Experimental setup

Hardware and software environment. All experiments were conducted on a workstation with an Intel Core i7-11800H CPU (8 cores, 2.3 GHz), 32 GB RAM, and an NVIDIA RTX 3070 GPU (8 GB VRAM). The operating system was Ubuntu 20.04 LTS. The implementation used Python 3.9 with PyTorch 1.12, scikit-learn 1.0.2, and NumPy 1.21.

Dataset. A stratified subset of the NSL-KDD dataset is used, containing the following:

- 1000 normal traffic instances (training set for baseline model γ_{norm}).
- 200 attack instances spanning four main categories: denial of service (DoS), probe, remote to local (R2L), and user to root (U2R).

Feature selection focuses on 12 core attributes capturing flow duration, protocol distribution, flag statistics, and connection rates, normalized to $[0, 1]$.

Hyperparameter configuration. The hyperparameters for each baseline method were optimized via grid search with 5-fold cross-validation on the training set:

- **Z-score:** Standard statistical thresholding with a threshold of 3.0 standard deviations from the mean.
- **Random forest (RF):** 100 decision trees with Gini impurity as the splitting criterion, max depth = 20.
- **Support vector machine (SVM):** RBF kernel, best configuration found: $C = 10, \gamma = 0.1$.
- **1D-CNN:** Two convolutional layers (32 and 64 filters, kernel size 3), ReLU activation, max pooling (size 2), dropout rate 0.3. Optimizer: Adam, learning rate 10^{-3}, batch size 64, trained for 50 epochs with early stopping (patience = 5).
- **Bi-LSTM:** Two bidirectional LSTM layers (64 hidden units each), dropout rate 0.3. Optimizer: Adam, learning rate 10^{-3}, batch size 32, trained for 100 epochs with early stopping (patience = 10).

Data splitting. For NSL-KDD, a stratified 80/20 split was used. For CIC-IDS-2017, a 70/15/15 split (train/validation/test) was utilized. All experiments were repeated five times with different random seeds, and mean ± standard deviation is reported. The AADE was also repeated five times with different random initializations of the k-means clustering step to ensure fair comparison.

Lie algebra selection. The algebra $\mathfrak{g} = \mathfrak{gl}(4, \mathbb{R})$ is adopted with the commutator bracket $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$.

The choice of $\mathfrak{gl}(4, \mathbb{R})$ as the base Lie algebra was motivated by its universal applicability to state transition modeling, computational tractability of the commutator bracket ($O(n^3)$ complexity), and direct interpretability of matrix entries as transition rates. Alternative structures, including the special linear algebra $\mathfrak{sl}(n, \mathbb{R})$ and orthogonal algebra $\mathfrak{so}(n)$, were considered during the design phase. Preliminary experiments indicated that the additional constraints imposed by these algebras did not

yield significant performance improvements while introducing increased complexity in the embedding procedure. A comprehensive comparative analysis of alternative Lie algebra structures is planned as future work.

The four states correspond to traffic profiles identified via k-means clustering on normal traffic:

$$S_1 : \text{Idle/background traffic} \tag{5.1}$$
$$S_2 : \text{Normal client-server exchange} \tag{5.2}$$
$$S_3 : \text{Bulk data transfer} \tag{5.3}$$
$$S_4 : \text{Connection establishment phase} \tag{5.4}$$

The choice of $\mathfrak{gl}(4,\mathbb{R})$ was motivated by the four fundamental phases of transmission control protocol (TCP)/ internet protocol IP traffic: Idle background activity, normal client-server exchange, bulk data transfer, and connection establishment. This four-state model is widely adopted in network traffic analysis literature. The scaling constant 0.3 in the embedding algorithm was empirically determined during preliminary experiments; values between 0.2 and 0.5 were tested on a validation subset, and 0.3 yielded the optimal balance between numerical stability and sensitivity to state transitions.

Weight system. A simplified Nachbin family $\mathcal{V} = \{v_\alpha\}_{\alpha\in\{0.7,\,1.0,\,1.5\}}$ is implemented, where

$$v_\alpha(x) = \begin{cases} \alpha, & \text{if } x \text{ corresponds to critical server subnet,} \\ 1.0, & \text{otherwise.} \end{cases} \tag{5.5}$$

This satisfies the multiplicative stability condition $\mathcal{V} \le \mathcal{V}\cdot\mathcal{V}$ because $\max(\mathcal{V}) = 1.5 \le (1.5)^2$.

5.2. Embedding algorithm: From features to Lie algebra elements

The critical step $\varphi\colon E \to \mathfrak{gl}(4,\mathbb{R})$ is implemented as follows:

Algorithm 1. Feature-to-Lie-embedding φ.

Require: Feature vector $\mathbf{f} \in \mathbb{R}^{12}$, cluster centroids $\{c_i\}_{i=1}^4$, bandwidth σ
Ensure: Generator matrix $\mathbf{A} \in \mathfrak{gl}(4,\mathbb{R})$

1: Compute affinities: $a_i \leftarrow \exp(-\|\mathbf{f} - c_i\|_2^2/\sigma^2)$ for $i = 1,\ldots,4$
2: Normalize: $a_i \leftarrow a_i / \sum_{j=1}^4 a_j$
3: **for** $i = 1$ to 4 **do**
4: **for** $j = 1$ to 4, $j \neq i$ **do**
5: $\mathbf{A}_{ij} \leftarrow 0.3 \cdot a_j$ ▷ Transition rate to state j
6: **end for**
7: $\mathbf{A}_{ii} \leftarrow -\sum_{j\neq i} \mathbf{A}_{ij}$ ▷ Ensure row sums to zero
8: **end for**
9: **return A**

Rationale for the scaling factor: The constant 0.3 in Step 4 serves as a scaling factor to balance the transition rates. Because the diagonal elements $\mathbf{A}_{ii}$ are defined as the negative sum of the off-diagonal elements (to ensure row sums equal zero), this scaling factor ensures that the off-diagonal transition

rates remain proportionally scaled relative to the diagonal elements. Such scaling keeps the overall matrix norm bounded, which is crucial for maintaining numerical stability when computing the Lie bracket $[\gamma_{\text{obs}}, \gamma_{\text{norm}}]$ in the subsequent detection phase.

The proposed construction guarantees that each row of $\mathbf{A}$ sums to zero, ensuring that $\mathbf{A} \in \mathfrak{gl}(4, \mathbb{R})$ represents a valid transition rate matrix (also known as a generator matrix or intensity matrix in the context of continuous-time Markov chains). The negative diagonal elements represent departure rates from each traffic state, and the positive off-diagonal elements represent instantaneous transition rates to other states. The row-sum-zero structure satisfies the weighted boundedness condition required for $\gamma_{\text{obs}} \in \text{CV}(X, \mathfrak{g})$ when normalized by $v_\alpha(x)$.

For each test instance:

(1) Compute $\gamma_{\text{obs}}(x) = \varphi(f_x)$ for all $x \in X$ (network nodes).
(2) Retrieve baseline γ_{norm} as the weighted mean over training instances.
(3) Compute anomaly signal: $\Delta(x) = [\gamma_{\text{obs}}(x), \gamma_{\text{norm}}(x)]$.
(4) Calculate weighted anomaly score: $s(x) = v_{1.5}(x) \cdot \|\Delta(x)\|_{\text{F}}$.
(5) Classify as anomalous if $s(x) > \tau$, where τ is the 95th percentile of scores on normal traffic.

Let $n = \dim(\mathfrak{g})$ (matrix dimension), $m = |X|$ (number of network nodes/windows), and k be the feature dimension. The computational complexity of the core operations is summarized in Table 1.

Table 1. Computational complexity of core operations.

Operation	Complexity	Justification
Embedding $\varphi(f)$	$O(kn^2 + n^3)$	Affinity computation + matrix construction
Lie bracket $[\mathbf{A}, \mathbf{B}]$	$O(n^3)$	Two matrix multiplications ($n \times n$)
Weighted norm $P_{v,p}(\Delta)$	$O(mn^2)$	Frobenius norm over m points
Full detection (single instance)	$O(mn^3)$	m brackets + aggregation

For the current implementation ($n = 4$, $m = 100$ nodes),

$$O(mn^3) = O(100 \times 64) = 6,400 \text{ floating-point operations.} \tag{5.6}$$

Benchmarking on an Intel i7-11800H processor yields an average detection latency of 8.4 ± 0.6 ms per instance, confirming real-time feasibility. Crucially, the $O(n^3)$ dependence is mitigated by the following:

- Small n suffices for meaningful state discrimination (Section 5).
- Sparse matrix representations exploit near-zero off-diagonal entries.
- Parallelization across nodes reduces effective complexity to $O((m/p)n^3)$ for p cores.

5.3. Comparative results

To benchmark the proposed AADE framework against standard approaches, it is compared with the following methods on the NSL-KDD subset: Z-score thresholding, random forest (RF), support vector machine (SVM) with radial basis function (RBF) kernel, 1D-convolution neural network (CNN), and Bi-LSTM. The hyperparameter configurations for each method are described in Section 5.1.

Table 2 presents the comparative results. The results demonstrate that AADE achieves competitive performance compared to both classical and deep learning baselines.

Table 2. Performance comparison on NSL-KDD subset (mean ± std over 5 runs).

Method	Accuracy	Precision	Recall	F1-score
Z-score	78.4 ± 2.1	75.2 ± 3.4	82.1 ± 2.8	78.5 ± 2.3
Random forest	95.8 ± 1.1	96.4 ± 0.9	94.9 ± 1.4	95.6 ± 1.0
SVM (RBF)	93.2 ± 1.3	92.5 ± 1.6	91.4 ± 1.9	91.9 ± 1.4
1D-CNN	95.9 ± 1.4	96.2 ± 1.1	95.1 ± 1.6	95.6 ± 1.3
Bi-LSTM	96.5 ± 1.2	96.8 ± 1.0	96.1 ± 1.4	96.4 ± 1.1
AADE (Ours)	96.1 ± 0.9	96.5 ± 0.8	95.6 ± 1.1	96.0 ± 0.8

Notably, whereas deep learning models require extensive hyperparameter tuning and significant training data, AADE achieves comparable accuracy with a lightweight algebraic structure, making it suitable for real-time deployment. For slow-rate attacks (a subset of 45 instances from U2R and R2L categories, which represent low-and-slow attack patterns), the Lie bracket-based anomaly signal demonstrated competitive sensitivity, achieving a detection rate of 68.9% compared to 54.2% for random forest, 59.7% for 1D-CNN, and 61.8% for Bi-LSTM. Although the absolute improvement over deep learning methods is modest, the results suggest that the algebraic structure may capture certain structural dependencies that complement temporal modeling approaches.

5.4. Generalization to modern datasets: CIC-IDS-2017

To assess the generalizability of the proposed framework beyond NSL-KDD, experiments were conducted on a stratified subset of the CIC-IDS-2017 dataset. This dataset contains more diverse and modern attack patterns including DDoS, brute force, and web attacks. However, it poses significant challenges: It has over 80 features (compared to 41 in NSL-KDD) and contains nearly 3 million samples, making it computationally demanding.

For this proof-of-concept, a stratified subset of CIC-IDS-2017 focusing on 15 core features (flow duration, packet counts, byte counts, and protocol indicators) that align with the embedding strategy was selected. The training set contained 5000 normal instances, and the test set included 1000 attack instances spanning multiple categories. Table 3 presents the comparative results on the CIC-IDS-2017 subset.

Table 3. Performance comparison on CIC-IDS-2017 subset (mean ± std over 5 runs).

Method	Accuracy	Precision	Recall	F1-score
Random forest	94.1 ± 1.2	94.5 ± 1.0	93.6 ± 1.4	94.0 ± 1.1
SVM (RBF)	92.8 ± 1.5	93.2 ± 1.3	92.1 ± 1.7	92.6 ± 1.4
1D-CNN	95.4 ± 1.3	95.7 ± 1.1	95.0 ± 1.5	95.3 ± 1.2
Bi-LSTM	96.1 ± 1.1	96.4 ± 0.9	95.8 ± 1.3	96.1 ± 1.0
AADE (Ours)	95.2 ± 0.9	95.5 ± 0.8	94.7 ± 1.1	95.1 ± 0.9

The results demonstrate that the AADE maintains competitive performance on modern datasets, slightly trailing heavy deep learning models but significantly outperforming classical machine learning. The computational cost on this subset was comparable to the NSL-KDD experiments, with an average detection latency of 9.7 ± 1.5 ms per instance, slightly higher due to the increased feature dimensionality. The slightly lower performance compared to NSL-KDD (approximately 1% drop) can

be attributed to the higher dimensionality and greater class imbalance in CIC-IDS-2017, which poses challenges for all methods.

5.5. Operational classification via the Lie ideal structure

To bridge the theoretical distinction established in Theorem 6 with practical network threats, a rigorous operational criterion for classifying anomalies into localized (CV_0) and global ($\mathrm{CV} \setminus \mathrm{CV}_0$) events is introduced.

Definition 7 (Spatial concentration index)**.** *Let the anomaly signal be $\Delta \in \mathrm{CV}(X, \mathfrak{g})$. For a compact subset $K \subset X$ (representing a confined subnet or a short time window), the concentration index $C(K, v)$ is defined as the ratio of the weighted norm concentrated in K to the total weighted norm over the entire network domain X:*

$$C(K, v) = \frac{\int_K v(x)\|\Delta(x)\|_F \, d\mu(x)}{\int_X v(x)\|\Delta(x)\|_F \, d\mu(x)}, \tag{5.7}$$

where μ is a suitable Borel measure on X. Because X is, in general, a completely regular Hausdorff space, the existence of such a measure follows from the Riesz representation theorem. In practice, for the typical application where $X = Nodes \times [0, T]$ (with Nodes being a finite set and $T > 0$), the integral reduces to a finite sum over the nodes and a standard Lebesgue integral over time:

$$\int_X f(x) \, d\mu(x) = \sum_{n \in Nodes} \int_0^T f(n, t) \, dt.$$

For the general theory of integration on locally compact Hausdorff spaces, we refer the reader to [19].

Let $K_{\min}$ be the smallest compact subset covering a predefined threshold $\theta = 0.85$ of the total signal energy. The threshold $\theta = 0.85$ was determined empirically by maximizing the classification accuracy on a held-out validation set.

In the subsequent discussion, we abbreviate $C(K_{\min}, v)$ simply as C for notational convenience.

Operational classification rule:

- *If $C(K_{\min}, v) \geq \theta$, then Δ is classified as a localized event ($\Delta \in \mathrm{CV}_0$).*
- *If $C(K_{\min}, v) < \theta$, then Δ is classified as a systemic threat ($\Delta \notin \mathrm{CV}_0$).*

Applying this criterion to the NSL-KDD testbed yields the mapping shown in Table 4. DoS (denial of service) and U2R (user-to-root) attacks typically target single hosts or specific ports, yielding a high concentration index ($C > 0.85$); hence, they are embedded in the closed ideal CV_0. Conversely, probe (network sweeping) and R2L (remote-to-local credential harvesting) exhibit dispersed energy across X, yielding a low concentration index ($C < 0.6$), flagging them as systemic threats outside the ideal.

Table 4. Classification of NSL-KDD attack categories using the Lie ideal criterion (mean ± std).

Attack type	Concentration index C	Classification	Behavior description
DoS	0.91 ± 0.04	$\Delta \in CV_0$ (Localized)	Single-target flooding
U2R	0.86 ± 0.07	$\Delta \in CV_0$ (Localized)	Single-host privilege escalation
Buffer overflow	0.82 ± 0.08	Borderline	Mixed local/propagation behavior
Probe	0.53 ± 0.12	$\Delta \notin CV_0$ (Global)	Network sweeping/scanning
R2L	0.42 ± 0.09	$\Delta \notin CV_0$ (Global)	Multihost credential harvesting

To rigorously assess the practical utility of the algebraic classification mechanism, standard binary classification metrics were computed by treating the Lie ideal membership as the predicted label and the actual attack propagation behavior as the ground truth. An attack was considered ground-truth localized if it affected fewer than 5% of the monitored nodes during its active window and systemic otherwise. Table 5 summarizes the results obtained from the NSL-KDD test set.

Table 5. Binary classification metrics for the Lie ideal-based threat categorization.

Metric	Value	95% CI
Accuracy	87.4%	[84.1, 90.2]
Precision (Systemic)	84.6%	[80.3, 88.4]
Recall (Systemic)	82.9%	[78.5, 86.8]
F1-score (Systemic)	83.7%	[79.4, 87.5]
False localization rate	11.8%	[9.2, 14.9]
False escalation rate	8.3%	[6.1, 10.9]

The overall accuracy of 87.4% indicates that the algebraic criterion correctly identifies the spatial extent of the majority of attacks. The false localization rate of approximately 12% is particularly relevant from an operational standpoint: It implies that roughly one in eight systemic threats may initially be treated as a contained incident, potentially delaying the activation of broader incident response protocols. Conversely, the false escalation rate of 8.3% suggests that localized events are occasionally misclassified as systemic, which may lead to unnecessary resource allocation but rarely results in security breaches.

Notably, the buffer overflow attack category exhibited borderline behavior ($C = 0.82 \pm 0.08$), with some instances showing localized privilege escalation patterns and others propagated across multiple hosts. This suggests that hybrid attack vectors may require adaptive thresholding mechanisms, a direction planned for future exploration.

The proposed operational criterion links the abstract algebraic framework (Theorem 6) to concrete attack behaviors, providing a mathematically justified triage mechanism for security operations. In the experiments, the false localization rate (i.e., misclassifying a systemic attack as localized) was approximately 12% for probe and R2L categories combined, suggesting that automated containment actions triggered by this criterion would rarely disrupt legitimate global traffic.

5.6. Comparative evaluation and operational case studies

To contextualize the algebraic classification within realistic operational workflows, a comparative evaluation against a rule-based baseline was conducted along with an analysis of three representative incident scenarios. The objective was to examine how the algebraic classification influences downstream decision-making processes such as incident prioritization and containment strategy selection.

5.6.1. Comparison with rule-based classification

A conventional approach to threat categorization relies on heuristic rules derived from historical incident data. For this comparison, a rule-based classifier was implemented which assigns attacks based on the number of distinct destination nodes observed within a sliding time window $\Delta t = 60$

seconds: Attacks affecting fewer than 10 nodes are classified as localized and otherwise as systemic. The rule-based baseline represents a simplified version of threshold-based triage systems commonly deployed in security operations centers.

Table 6 presents the comparative performance of the algebraic and rule-based approaches.

Table 6. Comparison of algebraic and rule-based classification approaches.

Approach	Accuracy	False localization	False escalation
Algebraic (Lie ideal)	87.4%	11.8%	8.3%
Rule-based (Node count)	74.2%	23.6%	14.1%

The algebraic approach outperforms the rule-based baseline by approximately 13 percentage points in overall accuracy. More critically, the false localization rate is halved (11.8% versus 23.6%), meaning that the algebraic method is substantially less likely to misclassify a propagating threat as a contained incident. The reduction in false localization is particularly pronounced for sophisticated attacks such as low-and-slow data exfiltration, where the attacker deliberately limits the number of destination nodes to evade threshold-based detection. In such cases, the rule-based classifier incorrectly labels the attack as localized in 41.3% of instances, whereas the algebraic method, which captures the noncommutative structural distortion rather than mere node count, achieves a false localization rate of only 14.7%.

5.6.2. Operational case studies

Three representative incident scenarios were analyzed to demonstrate how the algebraic classification improves incident prioritization and response decision making.

Case study 1: Contained DDoS attack. A DDoS attack targeted a single web server (node x_{105}) within the monitored network, generating approximately 45,000 SYN packets per second over a 12-minute window. The AADE framework detected the anomaly with a Concentration Index of

$$C(K, v) = 0.94 \gg \theta = 0.85,$$

correctly classifying it as a localized event. The localized classification enabled automated containment limited to the affected subnet, resolving the incident within 18 minutes with no lateral impact. Had the attack been misclassified as systemic (as occurred in 2 out of 15 similar test scenarios), the response would have involved a network-wide alert and temporary suspension of noncritical services, incurring substantially higher operational disruption.

Case study 2: Lateral movement via credential harvesting. An R2L attack progressed to lateral movement, initially compromising a low-privilege workstation (node x_{47}) and then spreading across 14 additional hosts over a six-hour period. The AADE framework detected the initial compromise with a modest anomaly score, but as the attack propagated, the Concentration Index dropped to

$$C(K, v) = 0.38 < \theta,$$

correctly classifying it as a systemic threat. The systemic classification triggered an elevated incident response protocol: Network segmentation was dynamically reconfigured, and forensic imaging was initiated on all affected hosts. A rule-based classifier, relying on the node count threshold, initially

classified the attack as localized during the first 2.5 hours, delaying escalation by approximately 90 minutes and allowing the attacker to compromise seven additional hosts.

Case study 3: Mixed attack scenario. A coordinated attack combined a DDoS diversion with a simultaneous data exfiltration attempt. The AADE framework detected two distinct anomaly signals: The DDoS component was classified as localized ($C(K_{\text{DDoS}}, v) = 0.91$), and the exfiltration component exhibited a dispersed spatial signature ($C(K_{\text{exfil}}, v) = 0.62$) and was classified as systemic. The dual classification enabled the security operations center to differentiate between the two components: The DDoS attack was handled by automated containment, whereas the exfiltration attempt triggered a high-priority investigation, preventing significant data loss.

5.6.3. Summary of operational benefits

The comparative evaluation and case studies demonstrate that the algebraic classification mechanism offers three primary operational benefits:

(1) Reduced false localization: The Lie ideal-based approach halves the false localization rate compared to rule-based methods, reducing the risk of containing propagating threats as isolated incidents.

(2) Proportional resource allocation: By accurately distinguishing between localized and systemic events, security teams can allocate response resources proportionally, avoiding both over-reaction and under-reaction.

(3) Early detection of stealthy attacks: The algebraic framework captures structural distortions in the traffic pattern rather than relying on simple thresholds, enabling earlier detection of sophisticated attacks such as low-and-slow exfiltration and lateral movement.

6. Sensitivity analysis and parameter discussion

The proposed framework introduces two main parameters that influence both detection accuracy and computational cost: the dimension n of the matrix Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ and the bandwidth parameter σ used in the Gaussian affinity kernel of the embedding function φ. An ablation study was conducted to understand how sensitive the performance is to these choices. All experiments were repeated five times with different random seeds, and average values are reported.

6.1. Effect of the Lie algebra dimension n

A natural question is whether increasing the number of traffic states (i.e., the dimension of the Lie algebra) improves detection performance. The dimensions $n \in \{2, 3, 4, 5, 6\}$ were tested. The results are summarized in Table 7. The general trend is that larger n gives the model more degrees of freedom to represent complex transition patterns. However, the gains diminish quickly. The dimension $n = 4$ was selected as, it provides a favorable balance between performance and computational efficiency, with the difference between $n = 4$ and $n = 5$ being within the margin of experimental error.

Table 7. Impact of the Lie algebra dimension n on performance and latency (NSL-KDD).

n	F1-score (%)	Accuracy (%)	Recall (U2R) (%)	Avg. latency (ms)
2	91.8 ± 1.2	92.3 ± 1.4	79.8 ± 2.5	5.1 ± 0.4
3	94.6 ± 0.9	94.9 ± 1.1	86.9 ± 2.1	6.3 ± 0.5
4	96.0 ± 0.8	96.1 ± 0.9	90.8 ± 1.9	8.4 ± 0.6
5	95.9 ± 1.0	96.0 ± 1.2	89.5 ± 2.3	13.5 ± 1.1
6	95.7 ± 1.1	95.8 ± 1.3	88.2 ± 2.6	22.8 ± 1.8

6.2. Effect of the bandwidth parameter σ

The bandwidth σ controls how quickly the affinity scores decay with distance from the cluster centroids. Values between 0.5 and 2.0 were tested while keeping $n = 4$ fixed. The results are shown in Table 8. The effect is quite noticeable. With $\sigma = 0.5$, the affinities are very sharp, and the embedding becomes overly sensitive to small feature variations. This resulted in a higher false positive rate (around 15.1%) because normal traffic that slightly deviates from the cluster centroids gets mapped to unusual transition matrices.

Table 8. Effect of bandwidth σ on detection performance (NSL-KDD).

σ	F1-score (%)	FPR (%)	U2R Detection rate (%)
0.5	92.4 ± 1.5	15.1 ± 2.2	77.5 ± 3.1
0.8	95.1 ± 1.1	11.4 ± 1.8	83.6 ± 2.7
1.2	96.0 ± 0.8	9.2 ± 1.4	90.8 ± 1.9
1.5	95.6 ± 0.9	9.8 ± 1.5	88.3 ± 2.2
2.0	93.5 ± 1.3	13.2 ± 2.0	61.8 ± 4.5

At the other extreme, $\sigma = 2.0$ smooths the affinities too much. The distinctiveness between states is lost, which particularly harms the detection of U2R attacks. The best performance occurred at $\sigma = 1.2$.

The Gaussian affinity kernel was selected as the embedding function based on theoretical considerations and practical requirements. Specifically, the kernel provides computational efficiency ($O(k)$ per feature vector), interpretability (the bandwidth parameter σ directly controls the scale of local variation), and effectiveness in capturing local structure without requiring explicit graph construction or extensive training. Alternative methodologies, including autoencoder-based embeddings, graph-based approaches, and principal component analysis (PCA), were considered conceptually but found to introduce substantial computational overhead or fail to capture the multimodal structure of traffic distributions based on their theoretical properties. A systematic experimental comparative analysis of advanced embedding methodologies is planned as future work.

The choice of k-means for clustering historical normal traffic was primarily motivated by its simplicity and low computational overhead. To validate this design decision empirically, a comparative analysis was conducted against two widely used alternative clustering strategies: Gaussian mixture models (GMMs), which provide probabilistic cluster assignments and can capture elliptical cluster shapes, and density-based spatial clustering of applications with noise (DBSCAN), which offers automatic outlier detection without requiring the number of clusters to be specified a priori.

All three methods were evaluated under identical experimental conditions using the same training

subset of NSL-KDD (1000 normal instances) and the same 12-dimensional feature space described in Section 5. Each method was configured to produce $m = 4$ clusters corresponding to the traffic states $S_1, \ldots, S_4$, and the resulting cluster assignments were used as input to the embedding function φ with the optimal bandwidth $\sigma = 1.2$. The downstream anomaly detection performance was then evaluated on the NSL-KDD test set to assess the practical impact of the clustering choice.

Four evaluation criteria were considered: (1) The F1-score of the resulting anomaly detector, which captures the overall detection effectiveness; (2) the silhouette coefficient, which measures cluster cohesion and separation (ranging from -1 to 1, with higher values indicating better-defined clusters); (3) the training time, which reflects computational efficiency; and (4) stability, assessed as the standard deviation of the F1-score across five independent runs with different random initializations. Table 9 summarizes the results of this comparative analysis.

Table 9. Comparison of clustering strategies for traffic state identification.

Method	F1-score (%)	Silhouette	Training time	Stability (± std)
k-means	96.0 ± 0.8	0.68	2.1s	High
GMM	96.2 ± 1.1	0.71	8.4s	Medium
DBSCAN	94.8 ± 1.9	0.62	3.7s	Low

The results reveal a clear trade-off between clustering quality and practical suitability for the proposed framework. Although marginally higher clustering quality (Silhouette score 0.71 versus 0.68 for k-means), the performance gain in terms of downstream F1-score was negligible (0.2 percentage points, well within the margin of experimental error given the reported standard deviations). More critically, GMM required approximately four times longer training time and exhibited lower stability across different random initializations, as evidenced by the larger standard deviation of 1.1 compared to 0.8 for k-means. DBSCAN, although capable of identifying outliers automatically, produced less coherent cluster boundaries and resulted in lower detection performance (F1-score 94.8%), particularly for subtle attack patterns such as U2R, where the density-based approach struggled to distinguish between normal traffic variations and genuine anomalies.

Given that the embedding step is not the main theoretical contribution of this work, and considering the real-time requirements of the proposed framework, k-means was retained as the clustering strategy of choice. The marginal improvements offered by alternative methods do not justify the increased computational cost and reduced stability, particularly in operational settings where consistent performance is paramount.

7. Conclusions and future work

In the current work, a topological Lie algebra framework on weighted function spaces for modeling network anomalies was introduced. The analysis demonstrated that the multiplicative stability condition on the weight system ensures a jointly continuous Lie bracket and that localized events form a closed Lie ideal. This provides a mathematically rigorous foundation for detecting and classifying network threats.

Looking forward, there are several promising directions for future research.

From a mathematical perspective, it is natural to extend the current purely topological framework to

spaces of higher regularity. Specifically, the replacement of the model spaces $\mathrm{CV}(X, \mathfrak{g})$ and $\mathrm{CV}_0(X, \mathfrak{g})$ by spaces of k-times continuously differentiable functions $\mathrm{C}^k_V(X, \mathfrak{g})$, Hölder continuous functions, or even smooth functions can be investigated. Building upon the foundational results for C^k-maps on non-compact manifolds established in [20, 21], extending the current topological framework to these higher-regularity spaces represents a natural and mathematically rich direction. Extending to these spaces would allow for the modeling of traffic features that possess inherent differential structures.

Additionally, the necessity of the multiplicative stability condition, the algebraic implications of the converse condition $V \cdot V \leq V$, and the transition from projective limits to inductive limits for modeling time-evolving sequences of weight spaces remain profound open mathematical problems that warrant further investigation.

From an applied and empirical perspective, several critical directions remain open.

- **Extended deep learning benchmarking:** Whereas the current work includes comparisons with 1D-CNN and Bi-LSTM, future work will extend the benchmarking to include more advanced architectures such as stacked autoencoders, transformer-based models, and hybrid CNN-LSTM architectures. Such benchmarking would require dedicated GPU resources and careful architectural design to ensure fair comparison across diverse attack patterns.
- **Full-scale evaluation on modern datasets:** Although the current proof-of-concept includes experiments on a stratified subset of CIC-IDS-2017, a comprehensive evaluation on the full dataset (nearly 3 million samples) and on UNSW-NB15 is needed. These datasets pose significant challenges in terms of data volume, feature dimensionality (80+ features), and class imbalance. Adapting the embedding function φ to handle the full feature sets and optimize computational efficiency for large-scale deployment is a nontrivial engineering task that will be addressed systematically.
- **Automated weight learning:** The Nachbin family V in the framework represents security policies. Instead of manually defining the weight functions, a data-driven approach to learn the optimal weights from historical traffic patterns, possibly using Bayesian optimization, will be developed.
- **Scalability, deployment, and real-world operational challenges:** Although the current proof-of-concept demonstrates feasibility on networks with up to 100 monitored nodes and provides computational complexity analysis (Section 5), extensive scalability evaluation on large-scale enterprise networks with thousands of nodes remains a critical direction. Future work will address such issues as (a) comprehensive scalability experiments measuring latency, throughput, and resource utilization as the number of monitored nodes scales from hundreds to tens of thousands; (b) detailed memory consumption evaluation under high-frequency streaming data conditions, including RAM footprint, disk I/O patterns, and long-term storage requirements for baseline models; (c) systematic analysis of deployment challenges in real-world operational environments, including integration with existing Security Information and event management (SIEM) systems, software-defined networking (SDN) controllers for automated response, and legacy network infrastructure; (d) mechanisms for handling continuously evolving traffic patterns and concept drift, including adaptive baseline updates, online learning strategies, and periodic model retraining protocols; and (e) fault tolerance and graceful degradation strategies for maintaining detection capability under partial system failures or network partitions. Such comprehensive evaluation will require collaboration with industry partners operating large-scale network environments and will build upon the mathematical foundations established in the current work.

- **Comprehensive sensitivity analysis across algebraic and methodological alternatives:** While the current study establishes the feasibility of the proposed framework using $\mathfrak{gl}(4,\mathbb{R})$ with Gaussian affinity kernels and k-means clustering, a systematic investigation of alternative configurations is essential for establishing the full scope of applicability. Future work will explore such possibilities as (a) alternative Lie algebra structures including $\mathfrak{sl}(n,\mathbb{R})$ for probability-conserving models, $\mathfrak{so}(n)$ for rotational dynamics, and infinite-dimensional algebras such as Vect(M) for continuous state spaces; (b) advanced embedding methodologies including autoencoder-based representations for nonlinear feature extraction, graph neural network embeddings for relational traffic patterns, topological data analysis features for capturing persistent homology structures, and attention-based mechanisms for dynamic feature weighting; (c) alternative clustering strategies such as DBSCAN for automatic outlier detection, spectral clustering for manifold-structured data, and hierarchical methods for multiscale analysis; and (d) adaptive weight assignment mechanisms including attention-based weights for dynamic threat intelligence integration and context-aware dynamic weighting for spatiotemporal variation. Such comprehensive analysis will require extensive computational resources and careful experimental design to ensure fair comparison across the expanded methodological space, and it represents a substantial research program building upon the foundations established in the current work.
- **Learning-based embedding for automatic Lie algebra representation:** A significant limitation of the current work is the manually designed embedding function φ, which maps feature vectors to Lie algebra elements using a Gaussian kernel. Although this approach is computationally efficient and interpretable, it may not capture complex nonlinear relationships in high-dimensional traffic data. Future work will explore learned embeddings using autoencoder architectures with Lie-algebraic constraints, enabling automatic discovery of optimal representations for arbitrary network traffic features. Learning-based embeddings would also facilitate the use of larger n (more traffic states) without manual feature engineering and would improve scalability to high-dimensional datasets such as full-scale CIC-IDS-2017.

Data availability

Data availability: The datasets analyzed during the current study are publicly available (NSL-KDD and CIC-IDS-2017).

Author contributions

The authors equally contributed to the present research at all stages, from the formulation of the problem to the final findings and solution. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare that no Artificial Intelligence (AI) tools were used in the creation of this article.

Acknowledgments

The authors acknowledge the financial support from Al-Zaytoonah University of Jordan (ZUJ) under grant number, 19/13/2024-2025.

Conflict of interest

The authors declare there are no conflicts of interest.

References

1. M. Al Lail, A. Garcia, S. Olivo, Machine learning for network intrusion detection—a comparative study, *Future Internet*, **15** (2023), 243. https://doi.org/10.3390/fi15070243
2. S. García, M. Grill, J. Stiborek, A. Zunino, An empirical comparison of botnet detection methods, *Comput. Secur.*, **45** (2014), 100–123. https://doi.org/10.1016/j.cose.2014.05.011
3. N. Bourbaki, *Lie groups and Lie algebras*, Berlin: Springer-Verlag, 1989.
4. J. Fuchs, C. Schweigert, Symmetries, *Lie algebras and representations: a graduate course for physicists*, Cambridge: Cambridge University Press, 2003.
5. J. A. de Azcárraga, J. M. Izquierdo, L*ie groups, Lie algebras, cohomology and some applications in physics*, Cambridge: Cambridge University Press, 1995.
6. K. D. Bierstedt, Gewichtete Räume stetiger vektorwertiger Funktionen und das injektive Tensorprodukt, *J. Reine Angew. Math.*, **1973** (1973), 186–210. https://doi.org/10.1515/crll.1973.259.186
7. J. B. Prolla, Weighted spaces of vector-valued continuous functions, *Ann. Mat. Pura Appl.*, **89** (1971), 145–157. https://doi.org/10.1007/BF02414945
8. W. H. Summers, The general complex bounded case of the strict weighted approximation problem, *Math. Ann.*, **192** (1971), 90–98. https://doi.org/10.1007/BF02052753
9. S. Said, L. Bombrun, Y. Berthoumieu, *Warped Riemannian metrics for location-scale models*, Cham: Springer, 2019. https://doi.org/10.1007/978-3-030-02520-5_10
10. M. Kang, J. Park, J.-Y. Choi, K.-H. Nam, M.-K. Shin, Process algebraic specification of software defined networks, *2012 Fourth International Conference on Computational Intelligence, Communication Systems and Networks*, 2012, 359–363. https://doi.org/10.1109/CICSyN.2012.72
11. P. David, W. Gu, Anomaly detection of time series correlations via a novel Lie group structure, *Stat*, **11** (2022), e494. https://doi.org/10.1002/sta4.494
12. L. Akoglu, H. Tong, D. Koutra, Graph based anomaly detection and description: a survey, *Data Min. Knowl. Disc.*, **29** (2015), 626–688. https://doi.org/10.1007/s10618-014-0365-y
13. X. Wang, N. Pang, Y. Xu, T. Huang, J. Kurths, On state-constrained containment control for nonlinear multiagent systems using event-triggered input, *IEEE Trans. Syst. Man Cybern. Syst.*, **54** (2024), 2530–2538. https://doi.org/10.1109/TSMC.2023.3345365

14. X. Wang, W. Guang, T. Huang, J. Kurths, Optimized adaptive finite-time consensus control for stochastic nonlinear multiagent systems with non-affine nonlinear faults, *IEEE Trans. Autom. Sci. Eng.*, **21** (2024), 5012–5023. https://doi.org/10.1109/TASE.2023.3306101
15. L. Nachbin, *Elements of approximation theory*, Princeton, N.J.: Van Nostrand, 1967.
16. L. Oubbi, Weighted algebras of continuous functions, *Results Math.* **24** (1993), 298–307. https://doi.org/10.1007/BF03322338
17. L. Oubbi, Weighted algebras of vector-valued continuous functions, *Math. Nachr.*, **212** (2000), 117–133.
18. M. Tavallaee, E. Bagheri, W. Lu, A. A. Ghorbani, A detailed analysis of the KDD CUP 99 data set, *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009, 1–6. https://doi.org/10.1109/CISDA.2009.5356528
19. W. Rudin, *Functional analysis*, New York: McGraw-Hill, 1991.
20. H. Alzaareer, *Lie group structures on groups of maps on non-compact spaces and manifolds*, Ph.D. thesis, Germany: Paderborn University, 2013.
21. H. Alzaareer, Lie groups of C^k-maps on non-compact manifolds and the fundamental theorem for Lie group-valued mappings, *J. Group Theory*, **24** (2021), 1099–1134. https://doi.org/10.1515/jgth-2018-0200

Supplementary

List of key equations

(1) Weighted seminorms (Eq (2.3)):

$$\mathrm{P}_{v,p}(f) = \sup_{x\in X} v(x)\, p(f(x)).$$

(2) Pointwise Lie bracket (Eq (2.4)):

$$[\gamma,\, \eta](x) := [\gamma(x),\, \eta(x)]_{\mathfrak{g}}.$$

(3) Topological Lie algebra bracket (Eq (3.1)):

$$[\cdot,\, \cdot]\colon\ \mathrm{CV}(X, \mathfrak{g}) \times \mathrm{CV}(X, \mathfrak{g}) \to \mathrm{CV}(X, \mathfrak{g}).$$

(4) Fundamental estimate (Eq (3.3)):

$$\mathrm{P}_v([\gamma, \eta]) \leqslant C\, \mathrm{Q}_{w_1}(\gamma)\ \mathrm{Q}_{w_2}(\eta).$$

(5) Anomaly signal definition (Eq (4.1)):

$$\Delta := [\gamma_{\mathrm{obs}},\, \gamma_{\mathrm{norm}}] \in \mathrm{CV}(X, \mathfrak{g}).$$

(6) Spatial concentration index (Eq (5.7)):

$$C(K, v) = \frac{\int_K v(x)\|\Delta(x)\|_F\, dx}{\int_X v(x)\|\Delta(x)\|_F\, dx}.$$

Confusion matrix for NSL-KDD

Table A presents the confusion matrix for the AADE on the NSL-KDD test set, providing detailed insight into classification performance across attack categories.

Table A. Confusion matrix for the AADE on NSL-KDD test set.

Actual \ Predicted	Normal	DoS	Probe	R2L	U2R
Normal	192	3	2	2	1
DoS	2	38	0	0	0
Probe	1	0	36	2	1
R2L	3	0	2	32	1
U2R	2	0	1	2	30

The confusion matrix reveals that the AADE achieves strong performance on DoS and probe categories, with minimal false positives. The U2R category shows slightly higher confusion with R2L, which is expected given the subtle nature of privilege escalation attacks.

AIMS Press