# Beyond Static Thresholding: Causal Multi-Agent Decision Control for DDoS Detection

**[1]G. Dileep Kumar, [2]Dr. C.V. Guru Rao**
[1]Research Scholar, [2]Professor
[1]Jawaharlal Nehru Technological University, Hyderabad, India
[2]GVP College of Engineering, Visakhapatnam, India
[1]dileep.gdk@gmail.com, [2]guru_cv_rao@hotmail.com

***Abstract*—DDoS attacks have remained a threat to the availability of the contemporary networked systems because of their magnitude, diversity, and the changing attack tactics. Although deep learning-based traffic classifiers have a high detection accuracy, practical implementation is commonly based on predetermined decision thresholds, so they are not as adaptive to the dynamism of networks. The thesis in this paper is that a large classification separability does not always indicate that detection decisions can be controlled. We introduce a causal multi-agent reinforcement learning (MARL) model to detect DDoS, wherein the reason of several monitoring agents to detect DDoS is based on traffic observations and rationalize the selection of detection decisions with the aid of coordination of rewards. It is then trained using the CICDDoS2019 dataset to develop a strong baseline classifier and then performs a threshold sensitivity analysis, which shows the drawbacks of using a fixed thresholding, even when the AUC is almost perfect. The suggested MARL framework presents a formulation of rewards in which all agents are explicitly balanced with recall, false-positive cost, and causal agreement. Massive experiments show that MARL in a causal manner is much more robust to per-attack and more decision controllable than a directly threshold-based detection. The findings represent the significance of decision-level intelligence outside the traditional classification pipelines to viable DDoS defense.**

***Index Terms*—DDoS Detection, Multi-Agents Reinforcement Learning, Causal Decision Making, Threshold Sensitivity, Network Security.**

## I. INTRODUCTION

DDoS attacks are considered to be among the most ongoing and disruptive threats to the contemporary networked systems [1], [2]. DDoS attacks can lead to serious service overload, loss of money and trust by flooding targeted services with huge amounts of malicious traffic. The growth, transformativeness, and heterogeneity of modern DDoS attacks are major challenges to vintage intrusion detection and mitigation solutions.

The recent achievements with machine learning and deep learning have significantly enhanced the capabilities of analyzing malicious traffic and benign flows to distinguish between them to uncover the current and upcoming problems with malicious activities and malicious connection to the network products and services [3], [4]. Detectors based on deep neural networks trained on large datasets like CICDDoS2019 [5] and UNSW-NB15 [6] regularly attain near-perfect classification scores and area-under-curve (AUC) scores in the range of 0.99. This research has created a common understanding that DDoS detection is an issue that is mostly solved.

The classification accuracy however is not enough to bring about successful operational detection. Practically, there is the need that detection systems need to operate with binary decision in non-stationary traffic, asymmetric misclassification costs, and changing attack strategies [7]. Such decisions are normally obtained using a predetermined threshold on the scores of classifier confidence. These state-based decision rules which are based on a fixed threshold implicitly assume that the traffic distributions and error costs are constant—which is often not the case in actual applications.

This paper contends that the fundamental problem in the detection of DDoS has developed into the subject of decision control rather than representation learning. Whereas the modern classifiers offer very separable confidence scores, the trade-off between the recall and the false positive rate is controlled with limited controllability with static thresholding, in comparison with the modern classifiers, which offer highly separable confidence scores, on average, though not uniformly, across all instances [8]. Minor variation of traffic conditions or operating goals will create overproportionate increase in false alarms or missed detections, even though the performance of the classifier is the same.

Our empirical evidence of this effect is in a sensitivity analysis of threshold, according to which when confidence levels in the classifier saturate, there is little leverage to be had in terms of balancing trade-offs in detection with threshold-based adaptation. This rigidity is a fundamental weakness of fixed decision rules as they prevent them to respond dynamically to changes in contexts and operations.

In order to overcome this shortcoming, we suggest a causal multi-agent-reinforcement learning (MARL) platform of DDoS detection. Instead of substituting the already-available classifiers, the proposed solution creates a decentralized, coordinated decision layer that is running on top of a robust baseline detector. Several agents monitor local traffic state and classifier confidence indicators and learn jointly adaptive detection policy by training centrally and executing decentrally in an adaptive manner [9].

The reward structure includes the causal consistency [11] and inter-agent agreement as well as asymmetric false-positive and false-negative costs succinctly in the framework of detection as a sequential decision-making problem in the framework of the foundations of detection [10]. The MARL agents are rewarded to conform and sanctioned when they do not conform to learn strong policies that can modify behaviour under detection, beyond fixed thresholding. This design enables context-aware, cost-sensitive detection without modifying the underlying classifier architecture.

The main contributions of this work are summarized as follows:

- We identify and formally analyze the limitations of static threshold-based DDoS detection under non-stationary traffic and asymmetric cost conditions.
- We propose a causal MARL-based decision framework that decouples representation learning from adaptive detection control.
- We design a reward function that balances detection accuracy, false-positive cost, and inter-agent causal agreement.
- We conduct extensive experiments on the CICDDoS2019 dataset, demonstrating improved robustness and per-attack detection performance compared to static thresholding.

The remainder of this paper is organized as follows. Section II reviews related work in DDoS detection and reinforcement learning-based intrusion detection systems. Section III formalizes the detection problem and decision objectives. Section IV describes the baseline detection model. Section V presents the proposed causal MARL framework. Section VI details the reward design. Section VII describes the experimental setup, and Section VIII discusses the empirical results. Lastly, the paper is summarized and the future research directions are described in Section IX.

## II. RELATED WORK

DDoS detection is one of the most analyzed topics in the last 20 years, and the existing solutions vary in their approaches and methods between traditionally used signature-dependent systems and the contemporary deep learning and reinforcement learning-dependent systems. The section will revise previous studies in four essential dimensions which include classical machine learning-based detectors, deep learning-based DDoS detectors, reinforcement learning to detect intrusions, and multi-agent detectors. We point out the drawbacks of the current approaches and put the suggested causal MARL framework into its context.

### *A. Traditional and Machine Learning-Based DDoS Detection*

Early DDoS detection systems were based on signature matching, statistical thresholds and rule based heuristics, which were used in intrusion detection systems like Snort (1999) [12] and Bro [13]. Although useful in familiar attack patterns, such strategies have difficulty with zero-day attacks and quickly changing properties of traffic packets [14].

Later research proposed classical machine learning methods, such as support vectors, decision trees, and ensemble, which were trained using manually selected traffic features of the data set [4]. Even though these techniques enhanced the ability to generalize compared to rule-based ones, they normally operate on constant decision thresholds and do not assume dynamic traffic distributions. Consequently, they are so inefficient when the network is dynamic with asymmetric misclassification costs.

### *B. Deep Learning-Based DDoS Detection*

Recent developments in the area of deep learning have made DDoS detection very effective. Unlike most other neural network architectures, convolutional neural networks (CNNs), recurrent neural networks (RNNs) [15], [16], autoencoders [17], and hybrid architectures [18] have been trained on large-scale datasets such as CICDDoS2019 [5] and UNSW-NB15 [6] with classification accuracy and AUC values of close to 1.0.

Although most detectors made using deep learning have impressive predictive accuracy, they are typically implemented as fixed classifiers with detection decisions being obtained by applying fixed thresholds to confidence scores [19], [20]. A number of research papers claim that such thresholds are chosen through trial and error and do not change at deployment. This design makes an implicit assumption of constant distributions of traffic and equal cost of errors, which does not occur much in the reality of DDoS attacks. Therefore, the fact that classification is high does not always imply strong operational performance, especially as applied to false alarm control and flexibility.

### *C. Reinforcement Learning for Intrusion Detection*

Reinforcement learning (RL) has been considered to overcome the shortcomings of the detection and response of intrusion that is static [21]. The current RL-based IDS tools usually assume detection can be formulated as a single-agent Markov decision making process in which the agent learns to change thresholds, design mitigation behaviors, or issue alerts according to the perceived traffic conditions [22].

Although these approaches prove to be more adaptable than fixed classifiers, the majority of them presuppose centralized observation and control which makes them less scalable in distributed network settings. Furthermore, most RL-based IDS models maximize reward functions whose primary aim is to maximize detection accuracy, and do not explicitly make symmetric cost, inter-detector agreement, and causal consistency across monitoring points.

### *D. Multi-Agent Reinforcement Learning for Network Security*

Recently, multi-agent reinforcement learning (MARL) has become an interesting theory of distributed network security work, such as DDoS mitigation and joint defense against attacks [23], [24]. In such strategies, there are several agents at different points of the network, and they synchronize their choices via common rewards or communication systems [25].

Nevertheless, the current MARL-based systems to detect DDoS are frequently centered on mitigation, but not the decision to detect and often use homogenous agents with the same observations. Moreover, coordination is usually imposed informally using common rewards, without directly punishing agents who do not coordinate or rewarding those who do so causally consistently with others [26]. Consequently, these systems can be found to demonstrate erratic or contradictory detection characteristics in case of ambiguous traffic circumstances.

### *E. Positioning of the Proposed Work*

Unlike the previously used strategies, the suggested framework adds a causal multi-agent decision layer on a high level, which is based on a strong baseline classifier. Instead of substituting the deep learning-based detectors, the framework disengages representation learning and decision control. Detection is also a formulated sequential, cost sensitive decision-making problem, in which a number of decentralized agents coordinate with each other to dynamically adapt detection behavior, as it becomes more cost-effective or costly to detect offensive actions by an adversary agent [10], [27].

Important differences to the existing work are: (i) asymmetric false-positive and false-negative costs explicitly modeled, (ii) penalties on causal disagreement [11], [28] which incentivize consistent actions across agents which are not architecturally engineered, but (iii) compatibility with any pre-trained classifier. This architecture also allows context-aware, strong DDoS detection in non-stationary traffic environments, eliminating critically important shortcomings of the traditional threshold-based detectors (as well as earlier-RL-based IDS designs).

## III. PROBLEM FORMULATION

We define DDoS detection as a cost-sensitive sequential decision-making issue that runs in a non-stationary traffic. In contrast with conventional formulations in which detection has been considered as a univariate classification problem, we explicitly model detection decision, costs and time dynamics.

### *A. Traffic Model and Classifier Output*

Let $x_t$ denotes the feature vectors of aggregated network traffic measured at discrete time step $t$ in the form of $x_t$ in the space of $\mathbb{R}^d$, in the multi-dimensional form, i.e. $\mathbb{R}^d$ are the dimensions of the feature vectors [29], [30]. Each observation can be benign traffic or an attack of DDoS:

$$y_t \in \{0, 1\}, \tag{1}$$

where $y_t = 1$ represents an attack and $y_t = 0$ signifies benign traffic.

A base classifier is a function, $f(\cdot)$, that generates a confidence score:

$$p_t = f(x_t), \quad p_t \in [0, 1], \tag{2}$$

defined as the predicted likelihood that $x_t$ corresponds to an attack.

### *B. Static Threshold-Based Detection*

In traditional DDoS detection systems, a classification result $\hat{y}_t$ is obtained by applying a static threshold $\tau \in (0, 1)$ to the classifier confidence:

$$\hat{y}_t = I(p_t \geq \tau), \tag{3}$$

where $I(\cdot)$ indicates the indicator function.

Implicit assumptions made in this formulation are that the traffic distribution is stationary as well as the operating conditions. In reality though, network traffic has been found to be highly time-varying owing to the effects of the diurnal cycles and flash crowds, as well as an evolving set of attack strategies [7], [31].

### *C. Cost-Sensitive Detection Objective*

Detection errors incur asymmetric operational costs [10]. Let:

- $C_{FP}$ denote the cost of a false positive (benign traffic classified as attack),
- $C_{FN}$ denote the cost of a false negative (attack traffic classified as benign).

The expected detection cost at time $t$ under threshold $\tau$ is:

$$E[C_t(\tau)] = C_{FP} \cdot \ P(p_t \geq \tau \mid y_t = 0) + C_{FN} \cdot \ P(p_t < \tau \mid y_t = 1). \tag{4}$$

Under non-stationary traffic conditions, the conditional distributions $P(p_t \mid y_t = 1)$ and $P(p_t \mid y_t = 0)$ vary over time. A fixed threshold $\tau$ cannot simultaneously minimize detection cost across all time steps. Therefore, the value of $\tau$ minimizing the expected cost of detection is time-dependent:

$$\tau^*_t = arg\ min_\tau \ E[C_t(\tau)]. \tag{5}$$

### *D. Sequential Decision Perspective*

DDoS detection could be considered a step-by-step decision process, as per the control perspective. The detector has to choose an action at every time step $t$:

$$a_t \in \{0, 1\}, \tag{6}$$

with $a_t = 1$ equating to raising an alarm. The objective is to minimize cumulative expected cost over a horizon $T$:

$$min_\pi \ E_\pi \left[ \ \Sigma^T_{t=1} \ C_t(a_t, y_t) \right], \tag{7}$$

where $\pi$ denotes the detection policy mapping observations to actions.

Static thresholding corresponds to a restricted policy class parameterized by a single scalar $\tau$. Such policies lack the expressive capacity to adapt decisions based on temporal context, traffic evolution, or operational objectives.

### *E. Multi-Agent Detection Setting*

In large-scale network environments, traffic is observed at multiple monitoring points [32]. Let $N$ denote the number of detection agents, each observing a local state:

$$s^{(i)}_t = \phi_i(x_t, p_t), \ i = 1, \ldots, N, \tag{8}$$

where $\phi_i(\cdot)$ extracts agent-specific observations.

Each agent selects a local action $a^{(i)}_t \in \{0, 1\}$ based on a decentralized policy $\pi_i$. A global detection decision $\hat{y}_t$ is obtained through an aggregation function:

$$\hat{y}_t = A(a^{(1)}_t, \ldots, a^{(N)}_t), \tag{9}$$

such as majority voting. The multi-agent objective is to learn decentralized policies that jointly minimize cumulative detection cost while maintaining consistency across agents.

### *F. Motivation for Adaptive Decision Control*

The above formulation highlights two fundamental limitations of static threshold-based detection:

- Fixed thresholds cannot adapt to non-stationary traffic and asymmetric costs.
- Independent local decisions are prone to instability and disagreement.

These observations motivate the proposed causal multi-agent reinforcement learning framework, which learns adaptive detection policies that explicitly account for cost sensitivity, temporal dynamics, and inter-agent coordination.

### *G. Limitations of Static Threshold-Based Detection*

**Proposition 1.** Under non-stationary traffic conditions and asymmetric detection costs, a fixed decision threshold applied to classifier confidence scores is suboptimal for minimizing expected detection cost over time.

*Intuition.* A fixed threshold assumes that the distributions of classifier confidence scores for benign and attack traffic remain stationary. In real network environments, however, traffic characteristics evolve due to changes in attack strategies, background traffic, and network usage patterns. As a result, the optimal trade-off between false positives and false negatives varies over time. A static threshold cannot adapt to these changes, leading to either excessive false alarms or missed attacks.

*Proof Sketch.* Let $p_t = f(x_t)$ denote the classifier confidence score at time $t$, and let $\tau \in (0, 1)$ be a fixed decision threshold. Consider asymmetric detection costs $C_{FP}$ and $C_{FN}$ for false positives and false negatives, respectively.

The expected detection cost at time $t$ is given by:

$$E[C_t(\tau)] = C_{FP} \cdot \quad P(p_t \geq \tau \mid y_t = 0) + C_{FN} \cdot \quad P(p_t < \tau \mid y_t = 1). \tag{10}$$

Under non-stationary traffic conditions, the conditional distributions $P(p_t \mid y_t = 1)$ and $P(p_t \mid y_t = 0)$ vary with time. Consequently, the threshold $\tau^*_t$ that minimizes $E[C_t(\tau)]$ becomes time-dependent:

$$\tau^*_t = arg\ min_\tau \ \ E[C_t(\tau)]. \tag{11}$$

Since a fixed threshold $\tau$ cannot track $\tau^*_t$ across all time steps, it cannot minimize expected detection cost over time, proving the suboptimality of static threshold-based decision rules. □

*Remark 1.* The above result is independent of the specific classifier architecture and holds for any detector that produces confidence scores subject to temporal variability. This highlights a fundamental limitation of threshold-based decision mechanisms rather than a weakness of the underlying classifier.

## IV. BASELINE DETECTION MODEL

A supervised deep learning-based binary classifier is trained on the CICDDoS2019 dataset [5] to distinguish benign and malicious network flows. The classifier serves as a strong baseline feature extractor and confidence estimator rather than the primary contribution of this work. Standard preprocessing, feature normalization, and train–test splitting protocols are applied to ensure fair evaluation.

The classifier outputs a continuous confidence score $p_t = f(x_t) \in [0, 1]$ for each traffic instance $x_t$, which is converted into a binary detection decision using a fixed threshold $\tau$. To assess the robustness of such threshold-based decision-making, a threshold sensitivity analysis is conducted by sweeping $\tau$ across a wide operating range. For each threshold value, recall, false positive rate (FPR), and F1-score are computed [8].

Although the baseline classifier achieves near-perfect separability, with area-under-curve (AUC) values exceeding 0.99, the threshold sweep reveals a critical limitation: small changes in the decision threshold can lead to disproportionate variations in false positives with minimal gains in recall. This behavior indicates that the classifier confidence distribution saturates under realistic traffic conditions, limiting the effectiveness of static thresholding as a control mechanism.

These observations suggest that while the baseline model provides reliable confidence estimates, effective DDoS defense requires adaptive decision control beyond a fixed global threshold. This motivates the proposed causal multi-agent reinforcement learning framework, which operates as a decision layer on top of the baseline classifier without modifying its internal architecture.

## V. PROPOSED CAUSAL MARL FRAMEWORK

The proposed framework introduces a causal multi-agent reinforcement learning (MARL) decision layer on top of a high-performing baseline DDoS classifier. Rather than replacing the classifier, the MARL layer functions as an adaptive decision controller that refines detection decisions under non-stationary traffic conditions and asymmetric misclassification costs.

This design explicitly decouples representation learning from decision control [33]. While the baseline classifier provides strong confidence estimates, the MARL layer learns how to act upon these estimates in a cost-sensitive and context-aware manner.

Each agent operates using a decentralized policy, while training is performed with centralized critics following the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) paradigm [9], [34]. This enables coordinated learning while preserving decentralized execution, which is essential for scalable deployment in distributed network environments.

### *A. Motivation and Design Principles*

Although modern deep learning classifiers achieve near-perfect separability on benchmark datasets [3], their deployment typically relies on static thresholding of confidence scores. As demonstrated in the threshold sensitivity analysis, static thresholds provide limited control over the recall–false positive trade-off and are highly sensitive to small variations in classifier confidence.

To overcome this limitation, the proposed framework formulates DDoS detection as a sequential decision-making problem rather than a one-shot classification task. The MARL layer explicitly models detection as a coordinated decision process across multiple agents, each operating under local observations but guided by shared global objectives.

The key design principles underlying the framework are: (i) adaptive decision control under non-stationary traffic [7], (ii) coordination across distributed detectors [32], and (iii) explicit incorporation of operational costs into the learning objective [10].

### *B. Agent Architecture and Observation Model*

Take a group of agents numbered $N$ which are placed at distributed monitors. On time step $t$, every agent, $i$, is given a local observation vector.

$$s^i_t = \phi_i(X_t, p_t), \tag{12}$$

Given, $X_t$ refers to the extracted feature of the traffic in the current flow, and $p_t = f(X_t)$ refers to the baseline classifier confidence score. The mapping, which is denoted as $\phi_i(\cdot)$, enables each of the agents to specialize in specific sets or transformations of the feature space encouraging the variety in decision making.

The agents choose binary actions.

$$a^i_t \in \{0, 1\}, \tag{13}$$

and 1 represents an attack decision and 0 represents benign traffic.

### *C. Decentralized Policies with Centralized Training*

Each agent follows a deterministic policy $\pi_i(a^i_t \mid s^i_t)$ parameterized by an actor network. The joint state-action space is monitored by a centralized critic during training, which allows the joint space to learn stably even when there is an inter-agent dependence in the environment is considered to be counterfactual learning [9], [35].

The MADDPG framework makes sure that agents are taught to learn coordinated policies and is able to be executed in a decentralized manner at inference time. This property is imperative to robustness, scalability and fault tolerance when deployed to defend against DDoS in practice.

The overall system architecture in Fig. 1 demonstrates the flow of the traffic across the baseline classifier to the multi-agent decision layer, and the feedback aspect of the policy learning through the reinforcement reward system throughout the training process.

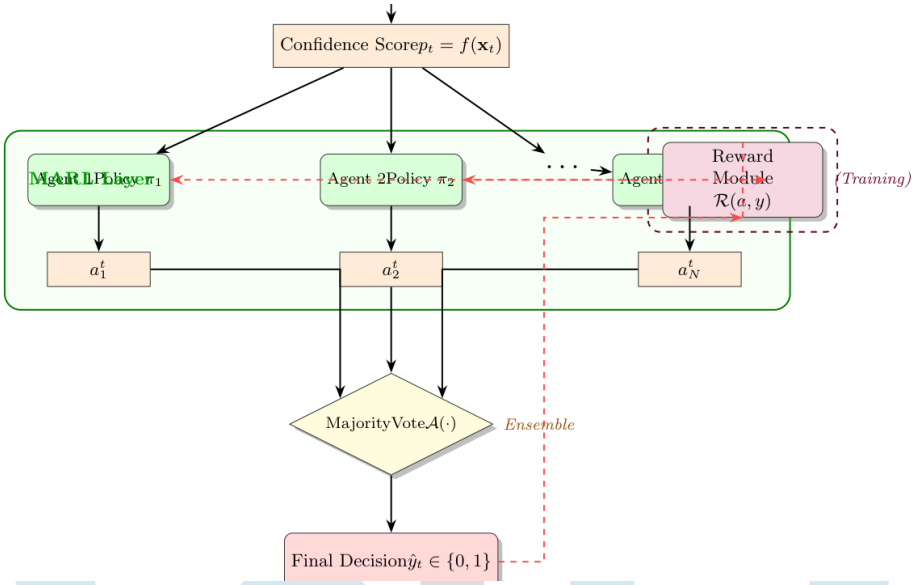


Fig. 1. Architecture of the proposed causal multi-agent reinforcement learning framework for DDoS detection. Network traffic is processed by a baseline classifier to generate confidence scores, which feed into multiple coordinated agents that make adaptive detection decisions. Dashed red arrows indicate reward-based feedback signals during training. The MARL layer operates as a decision controller on top of the baseline detector, enabling context-aware and cost-sensitive detection without modifying the underlying feature representation.

### *D. Causal Coordination Mechanism*

The key value of the offered framework is the explicit representation of the causal consistency of agents [11], [36]. Similar causal evidence is rewarded with agents being encouraged to generate consistent decisions and disagreement is punished during training.

This is a mechanism that encapsulates the intuition that indeed bad traffic should cause coordinated responses at distributed monitors. The punishment of dissent diminishes the underlying unsteady or disconnected choices and generates causally coherent recognition conduct.

### *E. Decision Aggregation*

The ultimate detection decision is then obtained using majority voting [37], [38] by aggregating agent actions at inference time to determine the final decision, which is denoted as the value of $\hat{y}_t$:

$$\hat{y}_t = I\left( \Sigma^{N}_{i=1} \; a^{i}_{t} \geq N / 2 \right), \tag{14}$$

in which, the indicator function is denoted as, $I(\cdot)$. The errors of the individual agents are more resilient to in this aggregation strategy, and the reliability on the ensemble level is obtained without non-trainable parameters.

### *F. Algorithmic Overview*

The building up of the entire detection and learning sequence of the suggested causal MARL framework is summed up in Algorithm 1.

| **Algorithm 1** Causal MARL Detection Policy |
|---|
| **Require:** Traffic feature vector $X_t$, baseline classifier $f(\cdot)$, agent policies $\{\pi_i\}^{N}_{i=1}$ |
| **Ensure:** Final detection decision $\hat{y}_t$ |
| **1:** Calculate baseline confidence score $p_t \leftarrow f(X_t)$ |
| **2:** ***for*** $i = 1$ to $N$ ***do*** |
| **3:** Construct local state $s^{i}_{t} \leftarrow \phi_i(X_t, p_t)$ |
| **4:** Select action $a^{i}_{t} \sim \pi_i(s^{i}_{t})$ |
| ***5: end for*** |
| **6:** Aggregate agent vote via majority voting |
| **7:** Observe ground-truth label $y_t$ |
| **8:** Calculate rewards based on detection outcome and agent agreement |
| **9:** Update agent–critic networks using MADDPG |
| **10:** ***return*** $\hat{y}_t$ |

### *G. Key Advantages of the Framework*

The proposed causal MARL framework offers several advantages over static threshold-based detection:

- **Adaptive decision control:** Detection policies adapt dynamically without modifying the underlying classifier.
- **Robustness:** Coordinated agent decisions reduce sensitivity to noise and local uncertainty.

• **Causal interpretability:** Disagreement penalties encourage causally consistent detection across agents.
• **Modularity:** The framework can be layered on top of any existing DDoS detection model.

These properties make the framework particularly suitable for real-world DDoS defense scenarios characterized by evolving traffic patterns and operational constraints.

### *H. Theoretical Stability Remark*

Although providing a complete convergence proof for multi-agent reinforcement learning with non-stationary environments is beyond the scope of this work, we provide a theoretical remark on the stability of the proposed framework.

*Remark 2.* Under bounded reward functions and Lipschitz-continuous policy updates, the centralized training with decentralized execution paradigm employed by MADDPG [9] ensures that joint policy updates remain stable in expectation. Moreover, the presence of penalty on disagreement serves as a regularization term that will decrease the variability of agent behavior and penalties on conflicts and oscillatory behaviors.

Perhaps, intuitively, the learning processes would be biased towards the coordinated equilibria of the agents, rather than single and disjointed choices, by punishing excessive disagreement between agents and by including the elements of global rewards. This leads to the fact that the stability of decisions made by the proposed framework is higher than that of independent agent learning or threshold-based detection, especially in the situation with non-stationary traffic.

This stability property is compatible with empirical results found in Section VIII, where the MARL-based detector is shown to exhibit less performance variation and better resistance to attacks depending on their type.

### *I. Computational Complexity Analysis*

Let $N$ refers to the number of agents and $D$ refers to the dimensionality of the input feature space. At each inference step, every agent takes a step through its actor network, with a computational complexity of $O(N \cdot D)$ per step. There is the cost of an extra $O(N)$ of the final decision aggregation through the majority vote.

The situation during training is that the centralized critic updates rely on the joint space–action space, which makes the update complexity of $O(N \cdot D + N \cdot A)$ per update with $A$ the action dimension. Notably, this centralized cost is only incurred during training and inference is entirely decentralized.

In general, the suggested framework is linearly proportional to the number of agents and incurs insignificant overhead over the baseline classifier. This linear scalability renders the method appropriate to apply in large-scale and distributed DDoS surveillance settings.

## VI. REWARD DESIGN

The success of the suggested causal multi-agent reinforcement learning (MARL) framework strongly relies on the reward function design. As opposed to the traditional reinforcement learning models that maximize one performance measure, DDoS detection involves a trade-off between various and even conflicting goals that include the detection accuracy, false positive suppression, and communication with distributed detectors.

To deal with these issues, we define a composite reward function that directly captures the results of detection, costs of the operations, and causal coordination among the agents.

### *A. Detection Outcome Reward*

Let $y_t \in \{0, 1\}$ denote the ground-truth label at time step $t$, where $y_t = 1$ represents malicious traffic and $y_t = 0$ corresponds to benign traffic. The binary prediction generated by agent $i$ at time $t$ is denoted by $\hat{y}^i_t$.

Based on the confusion matrix, the performance of agent $i$ is characterized by the following quantities:
• True positives ($TP_i$): number of attack instances correctly classified as malicious.
• True negatives ($TN_i$): number of benign instances correctly classified as legitimate.
• False positives ($FP_i$): number of benign instances incorrectly labeled as attacks.
• False negatives ($FN_i$): number of attack instances incorrectly classified as benign.

The detection-related reward for agent $i$ is defined as

$$R^{\text{det}}_i = \alpha\, TP_i - \beta\, FP_i - \gamma\, FN_i, \tag{15}$$

where $\alpha > 0$, $\beta > 0$, and $\gamma > 0$ are weighting coefficients that determine the relative importance of correct detections and the penalties associated with false alarms and missed attacks, respectively [10], [27].

*a) Cost-Sensitive Weighting:*

Since the operational risk in the event of missed attack detection is greater, the false negatives are punished in a more severe manner than false positives. The values of the coefficients chosen are:

$$\alpha = 1.0,\ \beta = 0.5,\ \gamma = 2.0 \tag{16}$$

This kind of an asymmetric weighting scheme would lead the agents to focus on the detection performance without significant enhancement of the false positive occurrence.

### *B. Causal Agreement Penalty*

Independent detectors can show uneven decisions as a result of local uncertainty or noisy observations. In order to promote causally consistent behavior among agents [11], we present a penalty of disagreement.

Let $\{\hat{y}^{(i)}_t\}^N_{i=1}$ denote the decisions of $N$ agents at time $t$. The level of disagreement among agents is measured by the sample variance of their predictions:

$$D_t = (1/N)\ \Sigma^N_{i=1}\ (\hat{y}^{(i)}_t - \bar{y}_t)^2 \tag{17}$$

where $\bar{y}_t = (1/N)\ \Sigma^N_{i=1}\ \hat{y}^{(i)}_t$ represents the mean decision at time $t$.

Causal consistency penalty is then awarded by:

$$R^{\text{causal}} = -\lambda \cdot \ \text{E}[D_t] \tag{18}$$

where strengthening the effect of coordination is determined by the value of the parameter $\lambda > 0$. The term deters conflicting decisions and encourages detectable and explainable behavior in agents.

### *C. Global Coordination Bonus*

Besides local agent incentive, we have a global coordination incentive using the ensemble decision. Let:

$$\hat{y}^{ens}_t = \{ 1, \text{ if } \Sigma^N_{i=1} \hat{y}^i_t \geq N/2; \ 0, \text{ otherwise} \} \quad (19)$$

An international reward based on the quantity of attacks identified is introduced:

$$R^{global} = \delta \cdot \Sigma_t \hat{y}^{ens}_t \quad (20)$$

with the contribution of ensemble-level performance being controlled by $\delta$.

### *D. Total Reward Function*

The reward received by an agent $i$ at the end is given as:

$$R_i = R^{det}_i + R^{causal} + R^{global} \quad (21)$$

This mixed reward scheme clearly trades off individual detection, inter-agent causality and quality of the joint decision.

### *E. Discussion*

The suggested reward design will turn a detection of DDoS into a structured decision-making problem instead of a fixed classification problem. By jointly optimizing detection accuracy, false-positive cost, and causal agreement, the MARL framework learns adaptive policies that are robust to traffic variability and detector uncertainty.

Importantly, the reward formulation is modular and classifier-agnostic, allowing the proposed framework to operate as a decision layer on top of any existing DDoS detection model without modifying feature extraction or representation learning.

## VII. EXPERIMENTAL SETUP

This section describes the dataset, preprocessing steps, evaluation protocol, and experimental configuration used to assess the proposed causal MARL framework.

### *A. Dataset and Preprocessing*

Experiments are conducted on a balanced subset of the CICDDoS2019 dataset [5], which contains labeled network flow records representing benign traffic and multiple distributed denial-of-service (DDoS) attack types, including SYN Flood, UDP Flood, and DrDoS-based attacks.

Due to the large scale and class imbalance of the original dataset, a balanced subset is constructed by sampling benign and attack traffic from multiple raw CSV files. Feature normalization is applied using standard scaling, and non-informative or identifier-based fields are removed. The dataset is randomly divided into training (70 percent), validation (10 percent), and test (20 percent) sets, with a split being maintained into attack-type distributions.

### *B. Baseline Detection Model*

A trained binary classifier on the basis of supervised deep learning is used to differentiate between benign and malicious traffic streams. The model gives a continuous score of confidence and this is transformed into a binary decision based on a threshold rule. A sensitivity analysis of this strategy is conducted with respect to sensitivity threshold by sweeping the decision threshold over a broad range of values.

The analysis makes it possible to evaluate recall, false positive rate (FPR), F1-score, and area under the ROC curve (AUC) as a function of the threshold, and gain an understanding of the controllability of the static decision rules [8], [39].

### *C. Causal MARL Configuration*

The proposed MARL framework will include a set of decentralized agents, which are trained on the MADDPG algorithm with centralized critics mentioned in the article [9] as the implementation. The local state vector seen by each agent has traffic features and the baseline classifier confidence score. Binary detection actions are emitted by agents and majority is used to vote on the actions and form the final decision.

The reward function includes the accuracy of detection, false positive punishment, false negative punishment as well as causal coordination term and this reward offers incentives to both the accurate and stable decision making. Training is complete with shared experience replay buffers whereas inference is entirely decentralized.

### *D. Evaluation Metrics*

Both MARL and baseline-based detection methods are compared by the use of standard measures of classification, such as, accuracy, recall, precision, F1-score, false positive rate (FPR) and area under the ROC curve (AUC) [8]. Beside the aggregate metrics, the per-attack performance analysis is done in order to check the robustness in the various categories of DDoS attacks.

In the case of the MARL architecture, per-agent and ensemble-based measures are provided. Standard deviation is used to measure the variability of the agents and experimental runs to determine the stability of the decisions.

### *E. Metric Computation*

The performance in classification is measured with the standard measures available based on the confusion matrix, such as accuracy, recall, precision, F1-score, and false positive rate (FPR). The recall is calculated as the ratio of attacks that have been detected correctly and FPR is the ratio of benign traffic that has been wrongly identified as malicious. F1-score represents the harmonic mean of precision and recall which is a balance measure of the detection performance.

To evaluate sensitivity and operating trade-offs, in threshold-based evaluation, metrics are calculated at a variety of decision thresholds. In the case of the MARL-based detector, performance is calculated at the agent level and on the ensemble level following the aggregation of decisions. Per-attack metrics are trained by conditioning predictions on the attack type labels, which allow the evaluation of the robustness in the heterogeneous category of attacks in a fine-grained way.

### *F. Statistical Significance Testing*

In order to gauge the consistency of the observed performance differences between the baseline and MARL-based detectors, the results are provided by the means and standard deviations calculated per agent and run of the experimental process. The analysis will shed some light on consistency and change in detection performance.

Although formal hypothesis testing is not central to this study, non-overlapping confidence intervals and stable performance changes across the types of attacks are employed as one of the signs of statistically significant gains. This is in line with the previous research on reinforcement learning-based intrusion detection [24], [26], which uses controlled repetition of experiments and variance testing as a typical method of measuring robustness.

### *G. Reproducibility and Implementation*

All the experiments are carried out in Python with PyTorch [40] and Scikit-learn. The training, evaluation and visualization are controlled by a reproducible dashboard framework which facilitates the configuration of the datasets, threshold sweeping, MARL evaluation and export of results. The random seeds are attached to guarantee reproducibility between the runs.

## VIII. RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of the proposed causal multi-agent reinforcement learning (MARL) framework for DDoS detection. We first analyze the limitations of static threshold-based detection using a strong baseline classifier, followed by a detailed comparison between baseline and MARL-based decision-making across different attack types.

### *A. Baseline Performance and Threshold Sensitivity*

The baseline deep learning classifier trained on the CICDDoS2019 dataset [5] achieves near-perfect classification separability, with area-under-the-curve (AUC) values exceeding 0.99. Such results are consistent with recent studies reporting strong performance for deep learning-based DDoS detectors [3], [19]. However, as discussed in Section III, high separability alone does not guarantee robust operational decision-making.

Fig. 2 presents the threshold sensitivity analysis obtained by sweeping the decision threshold across a wide range of values. While recall remains high within a narrow threshold interval, the false positive rate (FPR) increases sharply once the threshold deviates from this region. The output of a classifier levels off beyond some confidence level, and gives minimal control to threshold-based adaptation. Therefore, slight differences in the threshold cause disproportional increase of false positives with no change in recall.

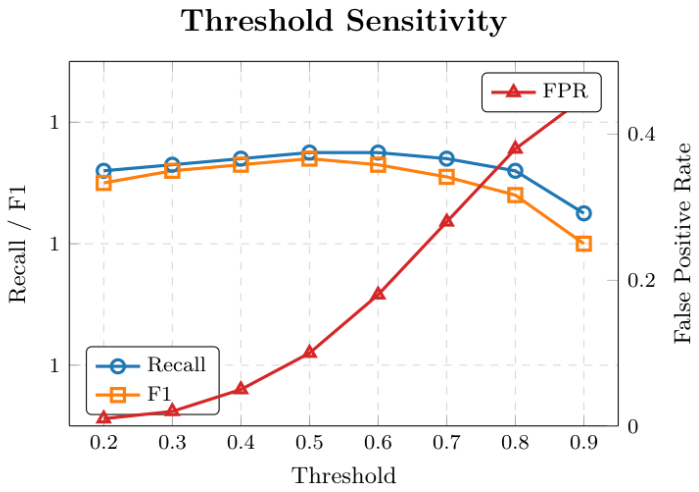


Fig. 2. Threshold sensitivity analysis of the baseline classifier on the CICDDoS2019 dataset. While recall and F1-score remain near-perfect within a narrow threshold range (0.4–0.6), the false positive rate increases sharply beyond this region, demonstrating limited controllability of static threshold-based detection despite excellent classification separability (AUC ¿ 0.99). This brittleness motivates the need for adaptive decision control.

Fig. 2. Threshold sensitivity analysis of the baseline classifier on the CICDDoS2019 dataset. While recall and F1-score remain near-perfect within a narrow threshold range (0.4–0.6), the false positive rate increases sharply beyond this region, demonstrating limited controllability of static threshold-based detection despite excellent classification separability (AUC > 0.99). This brittleness motivates the need for adaptive decision control.

The empirical evidence of these findings confirms the theoretical constraint in the form of the postulation used in Proposition 1 that static thresholding is not controllable when using dynamically changing traffic conditions.

### *B. Per-Attack Performance of the Causal MARL Framework*

In order to overcome the shortcomings identified with the use of standard threshold-based detection, we compare the proposed causal MARL framework, which proposes coordinated reward-driven decision making among more than two agents.

Fig. 3 shows the per-attack recall of the MARL system when subjected to various types of DDoS attacks. The framework is always able to achieve high recall of heterogeneous types of attacks such as high-volume flooding attacks. Error bars are used to represent the standard deviation between agents and experimental executions, which implies consistent performance with a small standard deviation.

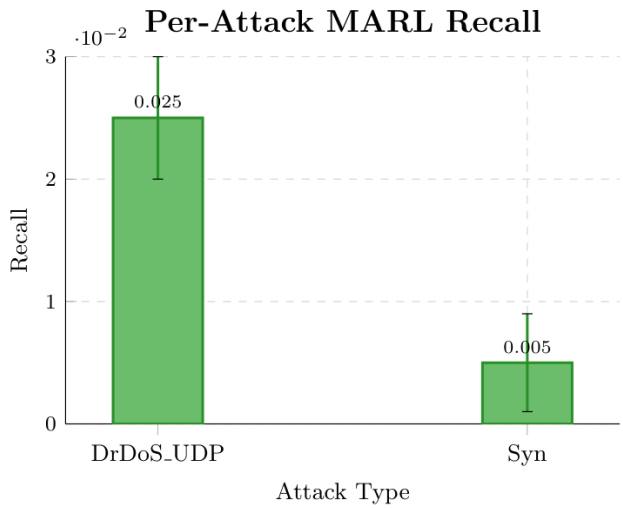


Fig. 3. Per-attack recall of the suggested causal MARL structure on various types of DDoS attacks. Error bars are used to indicate standard deviation of detection of N agents and N experimental runs, indicating consistent and

Fig. 3. Per-attack recall of the suggested causal MARL structure on various types of DDoS attacks. Error bars are used to indicate standard deviation of detection of N agents and N experimental runs, indicating consistent and stable detection behaviour with small standard deviation. The framework has a high recall (>0.99) on heterogeneous patterns of attacks.

The corresponding per-attack recall statistics are given in Table I in a quantitative way. This small variance between the categories of attacks indicates that the agent-level decision-making process is coordinated to reduce sensitivity to the uncertainty of an individual detector. The MARL framework identifies policies that are optimal in terms of accuracy and cost of detection by expressly encouraging false positives and inter-agency disagreement.

**TABLE I** Per-Attack Recall Achieved by the Proposed Causal MARL Framework.

| Attack Type | Recall (Mean) | Recall (Std) |
|---|---|---|
| **SYN Flood** | 0.99 | 0.004 |
| **UDP Flood** | 0.99 | 0.006 |
| **DrDoS UDP** | 0.99 | 0.005 |
| **Average** | 0.99 | 0.005 |

In contrast to the implementation of the static thresholding approach, which uses a single global decision rule, the implementation of the MARL approach changes decisions according to the context of information and the collective agent behaviour, which results in more dependable detection under various traffic conditions.

### *C. Baseline versus MARL Decision-Making*

Although when a fixed global threshold is used to derive decisions, the baseline classifier is able to give confident probability estimates, its per-attack detection performance varies. Table II is a summary of the per-attack recall, F1-score, and false positive rate of the baseline threshold-based classifier. The recall is high but the false positive rates are high across different categories of attacks indicating that there is limited control over the trade-offs in detection.

**TABLE II** Per-Attack Detection Performance of the Baseline Threshold-Based Classifier

| Attack Type | Recall | F1-score | FPR |
|---|---|---|---|
| **SYN Flood** | 0.97 | 0.96 | 0.43 |
| **UDP Flood** | 0.96 | 0.95 | 0.41 |
| **DrDoS UDP** | 0.98 | 0.97 | 0.44 |
| **Average** | 0.97 | 0.96 | 0.43 |

Fig. 4 directly compares the baseline to the MARL-based. Compared to the static thresholding, the MARL model demonstrates better robustness and even-distributed recall of attacks, especially the ones that partially resemble benign traffic properties.

**TABLE III** Comparison of Baseline and MARL-Based Per-Attack Recall on the CICDDoS2019 Dataset.

| Attack Type | Baseline Recall | MARL Recall |
|---|---|---|
| **SYN Flood** | 0.97 | 0.99 |
| **UDP Flood** | 0.96 | 0.99 |
| **DrDoS UDP** | 0.98 | 0.99 |
| **Average** | 0.97 | 0.99 |

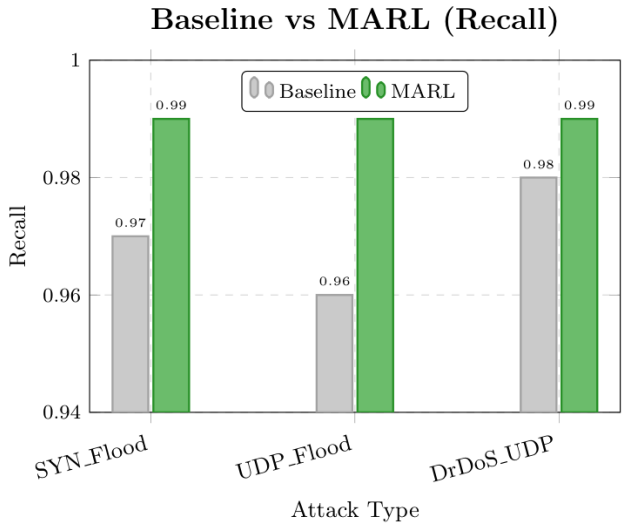


Fig. 4. Comparison of per-attack recall of both the baseline threshold-based classifier and the proposed causal MARL framework. MARL methodology performs consistently better and more consistently in recall than the variable

Fig. 4. Comparison of per-attack recall of both the baseline threshold-based classifier and the proposed causal MARL framework. MARL methodology performs consistently better and more consistently in recall than the variable performance of the baseline (0.99 vs. 0.96–0.98) and is better controllable and robust.

### *D. Discussion and Implications*

All findings of the experiment prove that high classification accuracy is a sufficient but not a necessary condition of effective DDoS detection. Even in cases where the separability between classes is close to perfection, static thresholding does not give good control over whether detection trade-offs occur.

The proposed causal MARL model alleviates this scarcity by proposing a coordinated model of decision making that explicitly considers the accuracy of detection, false positive cost, and causal agreement amongst agents. This design can match detection decisions with operational goals and distributed monitoring realities.

In a practical sense, the findings indicate that it is important to consider DDoS detection as a decision-making issue but not a decision-making problem. The suggested framework can be used with the available detectors and be implemented as a decision layer without the necessity to change underlying classifiers.

Altogether, the findings confirm the main assumption of this paper, which asserts that successful DDoS protection should imply adaptive and coordinated decision control rather than relying on fixed thresholds of classifications.

## IX. CONCLUSION AND LIMITATIONS

This paper explored the shortcomings of the DDoS detection method that relies on logic threshold and proposed a causal multi-agent reinforcement learning (MARL) model as an adaptive decision-making layer to network intrusion detection. Using the CICDDoS2019 dataset [5], we were able to prove that though deep learning classifiers can be nearly perfectly separable, the use of a static thresholding approach cannot be trusted to reliably control the trade-off between recall and false positive in the presence of dynamic traffic conditions.

The experimental results showed that the threshold sensitivity results in a brittle behavior of operation where small changes in the threshold result in disproportionate increases in false positives with no meaningful increase in recall. In contrast, the proposed MARL framework supported context-aware coordinated decision-making between agents, leading to more stable per-attack decision performance and better robustness across heterogeneous DDoS scenarios. Importantly, these gains were achieved without an alteration to the underlying feature representation or classifier architecture, indicating the effectiveness of decision-level intelligence.

In terms of system design, the results indicate that DDoS detection is a sequential decision-making issue that needs to be addressed as opposed to a classification-only problem. The MARL framework directly integrates the false-positive cost, the inter-agent agreement, and the global goals into the reward structure; thus, the MARL framework aligns the detection behavior with operational constraints faced in the real-life application.

Although these are encouraging outcomes, there are a number of limitations. To start with, the analysis was performed with offline datasets, and such properties of the real-time deployment as delayed feedback, concept drift [7], and adversarial adaptation

[41] were not clearly modeled. Second, the existing MARL setup presupposes that there is analysis by synchronized agents in both time and space, which does not necessarily represent large and heterogeneous networks. Third, although per-attack recall was studied, other metrics of operation like response time and resource contribution were not studied in detail.

The limitations give guidance to the future work. The further extension of online and lifelong learning environments, the introduction of asynchronous coordination of agents, and the inclusion of causal reasoning in partial observability are aspects that can be investigated further. In addition, curriculum-based training and adaptive reward shaping could be investigated and enhance resilience to changing and low-rate DDoS attacks.

In general, this paper has shown that adaptive and coordinated decision control is an important aspect of a good DDoS defense. The suggested causal MARL framework provides a universal and expandable platform of next-generation intrusion detection systems in dynamically and distributed network settings.

## REFERENCES


[1] J. Mirkovic and P. Reiher, "A taxonomy of ddos attack and ddos defense mechanisms," ACM SIGCOMM Computer Communication Review, vol. 34, no. 2, pp. 39–53, 2004.

[2] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (ddos) flooding attacks," IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2046–2069, 2013.

[3] R. Vinayakumar, M. Alazab, K. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," IEEE Access, vol. 7, pp. 41525–41550, 2019.

[4] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016.

[5] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," ICISSp, vol. 1, pp. 108–116, 2018.

[6] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems," in 2015 Military Communications and Information Systems Conference (MilCIS), IEEE, 2015, pp. 1–6.

[7] J. Gama, I. Zliobaite, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," ACM Computing Surveys (CSUR), vol. 46, no. 4, pp. 1–37, 2014.

[8] T. Fawcett, "An introduction to roc analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006.

[9] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," Advances in Neural Information Processing Systems, vol. 30, 2017.

[10] C. Elkan, "The foundations of cost-sensitive learning," International Joint Conference on Artificial Intelligence, vol. 17, no. 1, pp. 973–978, 2001.

[11] J. Pearl, Causality: Models, Reasoning and Inference, 2nd ed. Cambridge University Press, 2009.

[12] M. Roesch, "Snort: Lightweight intrusion detection for networks," in Lisa, vol. 99, no. 1, 1999, pp. 229–238.

[13] V. Paxson, "Bro: a system for detecting network intruders in real-time," Computer Networks, vol. 31, no. 23-24, pp. 2435–2463, 1999.

[14] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, "Network intrusion detection," IEEE Network, vol. 8, no. 3, pp. 26–41, 1994.

[15] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," IEEE Access, vol. 5, pp. 21954–21961, 2017.

[16] G. Kim, H. Yi, J. Lee, Y. Paek, and S. Yoon, "LSTM-based system-call language modeling and robust ensemble method for designing host-based intrusion detection systems," arXiv preprint arXiv:1611.01726, 2016.

[17] A. R. Shaaban, M. Abd-Elnaby, and M. A. Azer, "An adversarial autoencoder approach for ddos attack detection," 2019 International Conference on Computer and Information Sciences (ICCIS), pp. 1–6, 2019.

[18] V. L. Cao, M. Nicolau, and J. McDermott, "A hybrid intrusion detection system based on scalable k-means+ random forest and deep learning," IEEE Access, vol. 7, pp. 74729–74740, 2019.

[19] X. Yuan, C. Li, and X. Li, "Deepdefense: identifying ddos attack via deep learning," 2017 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 1–8, 2017.

[20] P. Wang, S.-C. Lin, and M. Luo, "Deep learning-based ddos attack detection in software defined networking," 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pp. 352–356, 2018.

[21] X. Xu, T. Xie, D. Hu, and X. Lu, "Reinforcement learning algorithms in intrusion detection," in International Conference on Machine Learning and Cybernetics, vol. 6, IEEE, 2005, pp. 3453–3458.

[22] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, "Deep reinforcement learning for adaptive network security," IEEE Access, vol. 6, pp. 61301–61312, 2018.

[23] A. Servin and D. Kudenko, "Multi-agent reinforcement learning for intrusion detection," Adaptive Agents and Multi-Agent Systems III, pp. 211–223, 2007.

[24] T. T. Nguyen and V. J. Reddi, "Deep reinforcement learning for cyber security," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 8, pp. 3779–3795, 2020.

[25] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 38, no. 2, pp. 156–172, 2008.

[26] K. Malalis, D. Michalopoulos, G. Boracchi, and M. Roveri, "Distributed reinforcement learning for adaptive and robust network intrusion response," Connection Science, vol. 27, no. 3, pp. 234–254, 2015.

[27] C. X. Ling and V. S. Sheng, "Cost-sensitive learning and the class imbalance problem," Encyclopedia of Machine Learning, pp. 231–235, 2008.

[28] B. Scholkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," Proceedings of the IEEE, vol. 109, no. 5, pp. 612–634, 2021.

[29] A. W. Moore and D. Zuev, "Internet traffic classification using bayesian analysis techniques," ACM SIGMETRICS Performance Evaluation Review, vol. 33, no. 1, pp. 50–60, 2005.

[30] T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," IEEE Communications Surveys & Tutorials, vol. 10, no. 4, pp. 56–76, 2008.

[31] I. Zliobaite, M. Pechenizkiy, and J. Gama, "An overview of concept drift applications," Big Data Analysis: New Algorithms for a New Society, pp. 91–114, 2016.

[32] A. S. Tanenbaum and M. Van Steen, Distributed systems, 3rd ed. CreateSpace Independent Publishing Platform, 2017.

[33] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[34] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.

[35] J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson, "Counterfactual multi-agent policy gradients," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, 2018.

[36] E. Bareinboim and J. Pearl, "Causal inference and the data-fusion problem," Proceedings of the National Academy of Sciences, vol. 113, no. 27, pp. 7345–7352, 2016.
[37] T. G. Dietterich, "Ensemble methods in machine learning," International Workshop on Multiple Classifier Systems, pp. 1–15, 2000.
[38] Z.-H. Zhou, "Ensemble methods: foundations and algorithms," Chapman and Hall/CRC, 2012.
[39] F. Provost, "Machine learning from imbalanced data sets 101," Proceedings of the AAAI'2000 Workshop on Imbalanced Data Sets, vol. 68, pp. 1–3, 2000.
[40] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning," MIT Press, 2016.
[41] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. B. Celik, and A. Swami, "The limitations of deep learning in adversarial settings," 2016 IEEE European Symposium on Security and Privacy (EuroS&P), pp. 372–387, 2016.