Research Annals of Industrial and Systems Engineering

www.raise.reapress.com

Res. Ann. Ind. Syst. Eng. Vol. 2, No. 4 (2025) 306–318.

Paper Type: Original Article

Game-Theoretic Analysis of Adaptive Intrusion Detection Systems under Detection Uncertainty

Mohammad Malekraeisi[1,*], Ali Akbar Kiani[1], Hamid Saadatfar[2]

[1] Department of Computer Engineering, Bi.C., Islamic Azad University, Birjand, Iran; mohammad.malekraeisi@iau.ac.ir.

[2] Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran; saadatfar@birjand.ac.ir

Citation:

Received: 20 July 2025 *Revised: 18 September 2025* *Accepted: 16 November 2025*	Malekraeisi, M., Kiani, A. A., & Saadatfar, H. (2025). Game-theoretic analysis of adaptive intrusion detection systems under detection uncertainty. *Research Annals of Industrial and Systems Engineering*, 2(4), 306-318.

Abstract

As computer networks expand and cyberattacks grow more sophisticated, Building Intrusion Detection Systems (IDS) that can withstand intelligent, adaptive adversaries has become a core problem in network security. Much of the existing literature assumes attacker behavior that is static or fully specified in advance, ignoring the uncertainty inherent to any real detection process, including false alarms and missed detections. Such simplifications make it difficult to capture how a defensive system and a learning attacker actually behave against each other over time. We address this gap with a game-theoretic framework, built on repeated games, that models the ongoing interaction between an IDS and an intelligent attacker. Both players are represented as autonomous agents that follow memory-one strategies and adjust their behavior through gradual learning. To make the model more realistic, detection uncertainty enters the security payoffs directly through the True Positive Rate (TPR) and the False Positive Rate (FPR), and a separate control parameter captures how quickly the attacker adapts relative to the defender, so we can study fast-adapting attacker scenarios. Numerical simulations produce a phase diagram over the space of attacker learning speed and IDS detection capability, revealing three distinct regimes: attacker dominance, adaptive competition, and a secure state. Faster-learning attackers push the security boundary outward, so the IDS needs stronger detection capability to stay in the secure regime. These findings offer practical guidance for designing and tuning adaptive IDS in dynamic, hostile network environments.

Keywords: Intrusion detection system, Network security, Adaptive attacker, Repeated games, Multi-agent learning.

1 | Introduction

The rapid growth of network infrastructures, cloud platforms, and Internet-based services has made network security central to stability and trust in modern information systems. Intrusion Detection Systems (IDS) sit at the front line of this defense, tasked with flagging malicious activity, cyberattacks, and unauthorized

Corresponding Author: mohammad.malekraeisi@iau.ac.ir

https://doi.org/10.22105/raise.vi.105

Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

behavior[1]. Yet traditional IDSs are increasingly strained by more sophisticated attacks, the rise of intelligent malware, and attackers who rely more and more on adaptive, learning-based techniques [2]. Much of the IDS literature still treats attacker behavior as static: fixed, predictable, or drawn from previously observed patterns [3]. Real-world attackers, however, increasingly use machine learning, watch how the defense responds, and gradually refine their strategy, probing IDS behavior and exploiting its blind spots [4]. In effect, they act as intelligent, adaptive agents continually optimizing their decisions inside a hostile environment. At the same time, no real IDS operates without uncertainty: false alarms and missed detections add operational cost and can themselves shape what an attacker does next [5]. Despite this, many analytical and simulation-based models either ignore this uncertainty or treat it too simply, pulling their results away from what actually happens on real networks [6]. Game theory has become a natural tool for capturing this strategic back-and-forth between defenders and attackers [7]. Rather than analyzing each side's decisions in isolation, security games treat them as part of one interactive process, with each move shaped by what the other player is doing [8]. Pairing game theory with multi-agent learning extends this further, giving researchers a framework for studying how strategies evolve, converge, and compete over the long run among intelligent agents [9]. Even so, three gaps stand out in the literature. Many game-theoretic security models still assume perfect, error-free detection. The role of the attacker's learning speed relative to the defender has rarely been studied systematically. And most work reports isolated, point-wise results rather than a regime analysis mapping out which regions of the parameter space are secure and which are not [10].

This paper takes these gaps as its starting point and proposes a repeated-game framework for the interaction between an adaptive IDS and an intelligent attacker. Both players are modeled as learning agents that follow memory-one strategies, and detection uncertainty is built directly into the payoff structure through the True Positive Rate (TPR) and the False Positive Rate (FPR). A separate parameter captures the attacker's learning speed relative to the defender, letting us explore a range of fast-adapting attacker scenarios. Extensive numerical simulations then yield a phase diagram over the key parameters, revealing three distinct regimes: attacker dominance, adaptive competition, and a secure state. Beyond the theoretical contribution, this regime analysis offers practical guidance for designing and tuning adaptive IDSs in dynamic network environments and for building defenses that hold up against intelligent attackers.

The rest of the paper is organized as follows. Section 2 surveys related work on IDS, security games, and learning-based approaches. Section 3 formulates the problem and lays out the game-theoretic framework connecting the IDS and the adaptive attacker, including the payoff structure that captures detection uncertainty. Section 4 turns to the learning dynamics, the memory-one strategies, and the role of the attacker's relative learning speed. Section 5 covers the experimental setup and evaluation procedure. Section 6 presents and discusses the results: strategy convergence, the regime phase diagram, and the influence of key parameters. Section 7 closes the paper with conclusions and directions for future work.

2 | Related Work

Prior work relevant to this study falls broadly into three strands: IDS and network-security modeling, game-theoretic approaches to cybersecurity, and adaptive multi-agent learning in attacker–defender settings.

2.1 | Overview of Intrusion Detection Systems and Game-Theoretic Models

IDSs are conventionally grouped into two families: signature-based and anomaly-based. Signature-based systems detect known attacks reliably but struggle against new or variant ones[1]. Anomaly-based systems take the opposite approach, modeling normal traffic so that deviations flag unknown attacks, at the cost of a higher false-alarm rate [11]. Machine learning has pushed detection accuracy considerably further, with neural networks, support vector machines, and deep-learning models all applied to this problem [3]. Even so, these approaches generally treat the IDS as a passive observer and largely overlook how an attacker's adaptive behavior affects the defense's long-term performance [12]. Many of them also focus narrowly on detection metrics, giving only limited attention to the operational costs that false alarms and detection errors actually impose [13]. Game theory occupies a distinct place in cybersecurity research because it models defenders and

attackers strategically rather than in isolation: each side's behavior is analyzed as a response to the other's strategy [6]. Security games of this kind have found wide application in critical-infrastructure protection, network security, and cyber defense more broadly [7]. Much of this work relies on static, complete-information games, where both players are assumed to know the full game structure and payoffs [14], an assumption that rarely survives contact with real networks. Repeated games have been proposed to capture longer-term interactions, but many still assume perfect, error-free detection [15], which again leaves little room to study how false alarms and detection errors actually shape attacker behavior.

2.2 | Adaptive Learning and Research Gap Analysis

Combining game theory with reinforcement learning and multi-agent learning to model dynamic security games is an active research area. Here, the defender and the attacker are both cast as learning agents that update their strategies based on environmental feedback [6], opening the door to studying convergence, instability, and adaptive competition directly. A number of studies look at adaptive attackers that alter their attack patterns to evade detection [2], but few examine systematically how the attacker's learning speed shapes the security boundary. As before, most of this work reports point-wise numerical results rather than a regime analysis that maps secure, competitive, and insecure regions across the parameter space [7].

Taken together, this body of work has advanced the use of game theory and adaptive learning in network security, but gaps remain: detection uncertainty is rarely built explicitly into game-theoretic models, the attacker's relative learning speed is seldom examined for its effect on security stability, and the lack of a comprehensive regime analysis makes it hard to pin down exactly when an IDS can hold a secure state against an intelligent attacker. The framework proposed here is aimed at closing these gaps, offering a systematic, multi-agent-learning-based analysis of the IDS-attacker interaction under realistic detection uncertainty.

3 | Problem Formulation and Game-Theoretic Framework

This section lays out the game-theoretic framework used to model the strategic interaction between the IDS and the adaptive attacker. The goal is to capture how both parties behave over the long run in a hostile, dynamic environment, accounting explicitly for detection uncertainty and each agent's capacity to learn.

3.1 | Security Game Structure and Payoffs

We model the IDS–attacker interaction as a two-player repeated game in which each side behaves as an autonomous, rational agent. Player one, the IDS, monitors the network and decides whether to activate its defensive mechanisms; player two, the attacker, chooses strategies to extract maximum benefit from its attacks. At every round of the game, the IDS chooses between:

I. Normal: Routine network monitoring, with no active defense engaged.

II. Active defense: Engaging advanced detection and defense mechanisms at some operational cost.

The attacker, in turn, chooses between:

I. No attack: Holding off.

II. Attack: Launching an attack on the network.

The game repeats over time, with each agent updating its decision based on the system's previous state and the feedback it receives; this is what allows the model to capture long-term, adaptive interaction [14]. The payoffs mirror the real costs and benefits faced by each side. A successful attack earns the attacker a benefit G and costs the defender a loss L. Engaging active defense costs the IDS an operational amount C_e, and mounting an attack costs the attacker C_a. When an attack is caught, the attacker pays a penalty P, and the IDS collects a matching reward. Together, these terms balance deterrence, operational cost, and potential attack payoff, a structure common throughout the security-game literature [6].

3.2 | Modeling Detection Uncertainty and Memory-One Strategies

Unlike much of the prior work, which treats detection as complete and error-free, this study builds the IDS's inherent uncertainty directly into the model through two parameters:

I. TPR: The probability that an attack is correctly identified when defense is active.

II. FPR: The probability of a false alarm when no attack has actually occurred.

When active defense is engaged, but no attack actually takes place, the IDS bears an extra cost, with probability equal to the FPR, standing in for wasted resources, service disruption, or the need for human intervention. When an attack does occur while defense is active, it is detected and blocked with probability TPR and succeeds with probability (1 − TPR). Framing the payoffs this way, in expectation, makes it possible to study the system's long-run behavior under realistically imperfect detection [11]. Both agents follow memory-one strategies: each decision depends only on what happened in the previous round, so the game's state space reduces to the four combinations of the IDS's and attacker's prior actions. Such a structure keeps the model tractable while still letting it capture reactive, adaptive behavior, an assumption widely used in the repeated-games literature [16].

3.3 | Learning Dynamics and Problem Formulation

Both agents update their behavior through gradual learning to maximize expected long-term payoff. To capture asymmetries in adaptation speed, we introduce a control parameter ω, defined as the ratio of the attacker's learning speed to the IDS's. Large ω values describe attackers that outpace the IDS in responding to environmental change, while small values reflect a defender with the learning advantage [17]. Put together, this study addresses the long-term behavior of a repeated security game played by learning agents, with payoffs shaped by detection uncertainty and asymmetric learning speeds: specifically, under what conditions the IDS can hold a secure state against an adaptive attacker, and how the security boundary moves across the parameter space.

4 | Learning Dynamics and Strategy Update Mechanism

This section describes how the IDS and the attacker gradually update their strategies. The framework rests on repeated games with memory-one strategies, in which each agent adjusts its policy parameters over time to maximize expected long-term payoff, a standard formulation in the repeated-games and multi-agent-learning literature for studying strategic, adaptive interaction [16], [17].

4.1 | Modeling Memory-One Strategies and Long-Term Payoff

As noted above, the system's state at step t is set by the combination of actions taken in the previous step, giving four possible states:

$$S = \{(N, NA), (N, A), (D, NA), (D, A)\}, \quad (1)$$

Here N and D denote normal monitoring and active defense, respectively, while NA and A denote no attack and attack. The IDS's memory-one policy is given by the probability vector

$$p = [p_1, p_2, p_3, p_4], \quad (2)$$

where $p_i = \Pr(D \mid s_i)$ is the probability of activating defense in state s_i, and the attacker's policy is defined analogously as

$$q = [q_1, q_2, q_3, q_4]. \quad (3)$$

with $q_i = \Pr(A \mid s_i)$. To keep dimensionality manageable while still reflecting realistic security behavior, we restrict the attacker's policy class: in the state offering a favorable, near-guaranteed opportunity, the attacker

attacks almost always, and in the high-risk state it almost never does. Concretely, two components of q are held fixed, and two are learnable, and symmetrically, only a subset of the IDS's parameters is learnable as well. Such a design keeps the learning process stable and makes the resulting analysis easier to interpret. Because the policies are memory-one, the game's dynamics reduce to a first-order Markov chain over the state space S: given p and q, the transition matrix M (p, q) follows directly, and the stationary distribution π satisfies

$$\pi = \pi M, \sum_{i=1}^{4} \pi_i = 1. \quad \textbf{(4)}$$

We compute π numerically by power iteration until it converges, and then obtain the expected long-term payoffs of the IDS and the attacker as expectations over the state distribution:

$$U_D(p,q) = \sum_{i=1}^{4} \pi_i \; r_D(s_i; p, q), \quad \textbf{(5)}$$

$$U_A(p,q) = \sum_{i=1}^{4} \pi_i \; r_A(s_i; p, q), \quad \textbf{(6)}$$

Here r_D and r_A denote the stage payoffs, which already incorporate imperfect detection through TPR and FPR along with the operational and false-alarm costs introduced in Section 3.

The IDS seeks to maximize U_D while the attacker seeks to maximize U_A, giving a two-objective game in which each agent nudges its policy parameters in the direction that improves its own long-term payoff:

$$\theta_D \rightarrow \theta_D + \alpha_D \nabla_{\theta_D} U_D, \quad \textbf{(7)}$$

$$\theta_A \rightarrow \theta_A + \alpha_A \nabla_{\theta_A} U_A. \quad \textbf{(8)}$$

Differentiating U analytically is awkward once the stationary distribution and the TPR/FPR-dependent payoffs are involved, so we approximate the gradient numerically via finite differences:

$$\frac{\partial U}{\partial \theta_k} \approx \frac{U(\theta_k + \varepsilon) - U(\theta_k - \varepsilon)}{2\varepsilon}, \quad \textbf{(9)}$$

where ε is a small step size used for the approximation.

4.2 | Asymmetric Learning Dynamics and Numerical Stability

To capture fast-adapting attackers, we set the attacker's learning rate relative to the IDS's as

$$\alpha_A = \omega \, \alpha_D, \omega \in (0,1], \quad \textbf{(10)}$$

As ω grows, the attacker reacts to change more quickly, and the competition intensifies. This single parameter drives the regime, or phase-diagram, analysis presented later.

To keep the numerics stable and the policies valid probabilities, we clip the parameters to [ε, 1 − ε] after every update:

$$p_i \rightarrow \text{clip}(p_i, \epsilon, 1-\epsilon), q_i \rightarrow \text{clip}(q_i, \epsilon, 1-\epsilon), \quad \textbf{(11)}$$

Clipping keeps probabilities from getting stuck at 0 or 1 and avoids numerical problems when computing the stationary distribution. The fixed parameters within the restricted policy class are similarly re-constrained at every iteration so that the intended policy structure is preserved.

Convergence is tracked by monitoring how much the stationary distribution, or the payoffs, change across successive iterations. In practice, the implementation runs a fixed number of iterations to reach steady state, computing the long-term payoff from π at each learning step.

4.3 | Proposed Learning Algorithm (Executive Summary)

The overall learning process is presented below in algorithmic form.

Algorithm 1. Gradual learning of the IDS and the attacker within a repeated-game framework with imperfect detection.

Initialize p and q

Enforce probabilistic constraints on p and q

for t = 1 to T do

Compute stationary distribution π for Markov process M (p, q)

Compute long-term payoffs:

U_D (p, q) → IDS payoff

U_A (p, q) → Attacker payoff

Approximate IDS gradient using finite differences:

$\nabla_p U_D$ (p, q)

Update IDS strategy:

$p \rightarrow p + \alpha_D \nabla_p U_D$

Approximate attacker gradient using finite differences:

$\nabla_q U_A$ (p, q)

Set attacker learning rate:

$\alpha_A = \omega \alpha_D$

Update attacker strategy:

$q \rightarrow q + \alpha_A \nabla_q U_A$

Apply clipping and policy constraints (if required)

end for

Set $p^* \rightarrow p$, $q^* \rightarrow q$

Return p^*, q^*, U_D (p^*, q^*), U_A (p^*, q^*)

4.4 | Direct Connection to Regime Analysis

Because ω and the TPR/FPR pair together govern the system's behavior, the game can settle into qualitatively different regimes depending on where these parameters sit. Running the algorithm above over a grid of (ω,

TPR) values produces a regime map delineating the attacker-dominance, adaptive-competition, and secure regions, giving a systematic way to determine when the IDS, despite operational costs and detection errors, can still maintain deterrence against a learning attacker.

5 | Experimental Setup and Performance Evaluation

5.1 | Simulation Configuration and Evaluation Metrics

To evaluate the model, we simulated the IDS–attacker interaction as a two-agent repeated game, with both sides following memory-one strategies so that each round's decision depends only on the previous round's state. Payoffs were computed in expectation, folding in TPR and FPR so that detection uncertainty is built directly into the simulation. The payoff parameters, including the benefit of a successful attack (G), the defender's loss on a breach (L), the cost of active defense (C_d), the cost of mounting an attack (C_a), and the detection penalty (P), were varied over reasonable ranges for the analysis. Adaptive behavior was modeled by updating each strategy's selection probability gradually, based on cumulative payoff, and the control parameter ω governed the attacker's learning speed relative to the defender, covering scenarios where the attacker learns more slowly, at the same pace, or faster than the IDS. Simulations ran over a long time horizon so that behavioral stability and payoff convergence could both be assessed. We evaluated performance using the two agents' expected long-term payoffs together with a regime analysis over the parameter space: extracting each agent's average steady-state payoff and mapping out the attacker-dominance, adaptive-competition, and secure regions, which together indicate what conditions are needed for stable security.

6 | Experimental Results and Analysis

This section presents and discusses the simulation results for the proposed IDS–attacker framework, focusing on the convergence of the learning dynamics, the effect of detection quality, the regime analysis in the (TPR, ω) space, and a joint look at security and the economic efficiency of defense.

6.1 | Convergence of the Learning Dynamics

We first checked the convergence behavior of both agents' policies and long-term payoffs to rule out any influence from transient learning effects on the final results. *Fig. 1* shows that the IDS's long-term payoff, across different ω values, settles into a steady state after an initial transient, evidence that the update mechanism, which relies on maximizing long-term payoff through the Markov chain's stationary distribution, is numerically stable. *Fig. 2* shows the same pattern for the attacker's long-term payoff, which converges without notable oscillation in the later stages of learning. Together, these results indicate that the two agents' adaptive interaction settles into a stable equilibrium, making meaningful long-term analysis possible.

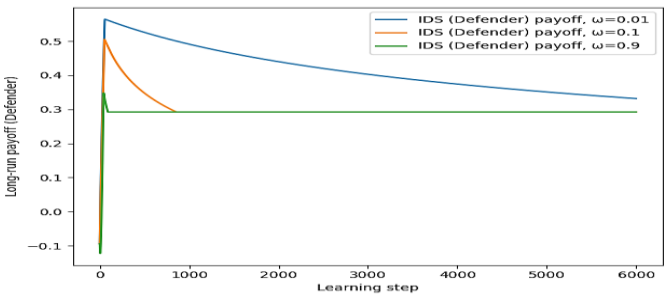


Fig. 1. Convergence of the IDS's policy parameters over learning iterations.

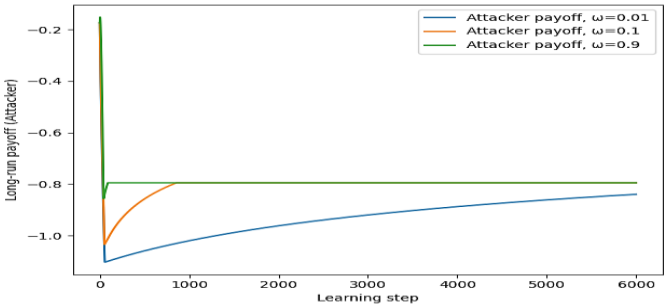


Fig. 2. Convergence of the attacker's policy parameters over learning iterations.

We also examined how the IDS's and the attacker's probabilistic policies converge. *Figs. 3* and *4* show that the action-selection probabilities across the different game states settle at fixed values, confirming that the learned policies are stable, further evidence that the framework consistently captures stable behavior for both learning agents.

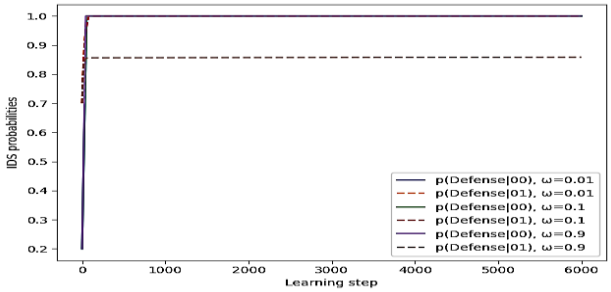


Fig. 3. Variation of the IDS's long-term payoff over iterations (convergence of U_D).

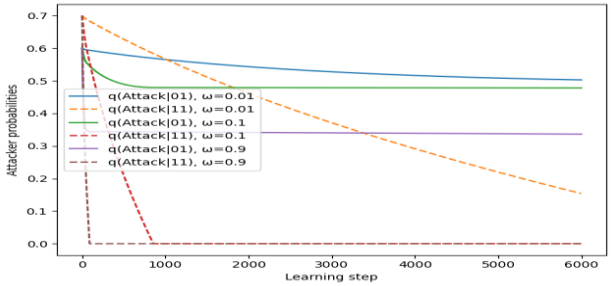


Fig. 4. Variation of the attacker's long-term payoff over iterations (convergence of U_A).

6.2 | Effect of Intrusion Detection Systems Quality

Here we examine how the IDS's detection quality shapes the game's strategic dynamics, through a sensitivity analysis over the TPR and the learning-speed parameter ω. *Table 1* reports only representative and boundary points, those marking a shift in behavioral regime, to keep the discussion focused. At low detection quality (TPR = 0.5, for instance), the IDS's expected payoff stays negative regardless of ω, while the attacker earns either a positive payoff or only a marginal loss. Such a pattern is a clear signature of attacker dominance: attacking remains strategically worthwhile, and active defense fails to deter. As the TPR rises further, the system passes through a transition region. Around TPR = 0.6 in *Table 1*, depending on ω, the regime shifts from attacker dominance into a mixed, arms-race-like state, where both players' payoffs are negative, yet neither has a clear incentive to disengage; strategic updates instead settle into an unstable, competitive equilibrium. This finding shows that improving detection quality alone does not guarantee security; the attacker's adaptive behavior still plays a decisive role. Once TPR crosses roughly 0.72, *Table 1* shows the regime shifting qualitatively into the secure region: the IDS's expected payoff turns positive while the

attacker's payoff drops substantially, the signature of stable deterrence, where attacking is no longer strategically attractive. The learning-speed parameter ω also matters for where these boundaries fall: at intermediate TPR values, *Table 1* shows that changing ω alone can push the system between the mixed and secure regimes. Security stability, in other words, depends not just on raw detection quality but also on how the IDS's learning process weighs past experience. Overall, these results suggest that network security is not simply a direct function of the TPR but emerges from the interplay between detection quality and adaptive learning dynamics. For IDS design, the practical takeaway is that stable security requires crossing specific thresholds on both the detection and the learning dimensions, not just improving one in isolation.

Table 1. Configuration of TPR/FPR test scenarios and summary of key results (average U_D, average U_A, and resulting regime).

Regime	Attacker payoff	IDS payoff	Learning-memory parameter (ω)	TPR
Attacker dominance	0.031	−0.371	0.01	0.5
Attacker dominance	−0.195	−0.193	0.01	0.6
Adaptive competition	−0.2	−0.191	0.42	0.6
Secure	−0.717	0.242	0.01	0.7
Adaptive competition	−0.661	0.19	0.17	0.7
Secure	−0.775	0.263	0.5	0.8
Secure	−1.015	0.476	0.5	1

6.3|Regime Analysis in the (TPR-ω) Space

To build a fuller, more systematic picture of the long-run IDS–attacker interaction, we carried out a regime analysis over the two-dimensional (TPR, ω) space, where ω captures the attacker's learning speed relative to the IDS, and TPR captures the IDS's intrinsic detection quality. This analysis lets us study the joint effect of the attacker's learning capability and the defense's detection accuracy on network-security stability. The system's behavior falls into three qualitative regimes: attacker-dominant, competitive (arms-race), and secure. Attacker dominance shows up mainly at low TPR, where the IDS cannot build effective deterrence and attacking still pays off economically; the defense fails against an adaptive attacker even though the IDS itself is learning. As detection quality improves, the system moves through a transition region that behaves like an arms race. Both agents keep updating their strategies; neither gains a lasting edge, and network security ends up hinging heavily on the learning parameters and operational conditions. *Table 2* quantifies these transition thresholds and shows that the TPR needed to enter the secure regime rises with ω. For small ω, a slowly adapting attacker, the IDS can secure the system at a fairly moderate TPR. As ω grows and the attacker reacts more quickly to past experience, the minimum TPR needed to stay secure climbs substantially. This coupling between ω and TPR is a key design insight: raising detection accuracy alone, without regard for the attacker's learning capability (or vice versa), does not guarantee stable security. Network security instead depends on a careful balance between detection quality and both players' learning dynamics. More broadly, the regime analysis here shows that the long-term IDS-attacker interaction can be mapped as distinct qualitative regions in the (TPR, ω) space, a framework that not only explains the underlying behavior but also yields practical thresholds for designing and tuning adaptive IDSs, linking game-theoretic analysis directly to network-security engineering.

Table 2. Sample boundary points: minimum TPR required to enter the secure regime for several values of ω.

Regime Above Threshold	Regime Below Threshold	Minimum TPR For Secure Regime	Ω
Secure	Attacker dominance	≈ 0.72	0.01
Secure	Attacker dominance	≈ 0.72	0.05
Secure	Adaptive competition	≈ 0.74	0.1
Secure	Adaptive competition	≈ 0.78	0.3
Secure	Adaptive competition	≈ 0.80	0.5
Secure	Adaptive competition	≈ 0.82	0.7
Secure	Adaptive competition	≈ 0.88	0.9

6.4 | Long-Term Payoff Map and Practical Interpretation of Results

Complementing the regime analysis above, this section looks at the IDS's long-term payoff map over the (TPR, ω) space, which lets us weigh security stability against the economic efficiency of defense together, going beyond the purely regime-based view. The heat map in *Fig. 6* shows the IDS's payoff generally improving as TPR rises; at low TPR, the payoff is mostly negative, reflecting an ineffective defense against an adaptive attacker. These low-payoff regions largely coincide with the attacker-dominant regime in *Fig. 5*; under these conditions, the IDS is doing poorly both from a security standpoint and a cost-benefit one. At intermediate TPR (matching the mixed/arms-race regime in *Fig. 5*), the IDS's payoff hovers near zero or turns slightly positive, a borderline situation in which defense and attack are roughly balanced and small changes in learning parameters or operational cost can tip the advantage either way. Practically speaking, this is where network security is most fragile and where IDS parameters need the most careful tuning.

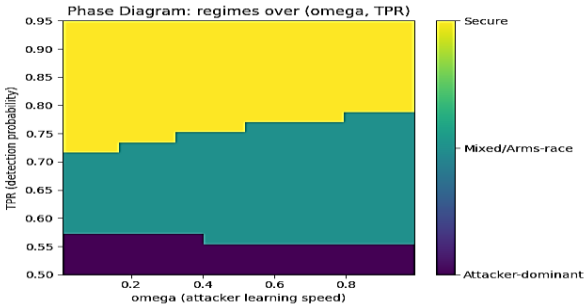


Fig. 5. Phase diagram in the (TPR-ω) space showing the three regime regions.

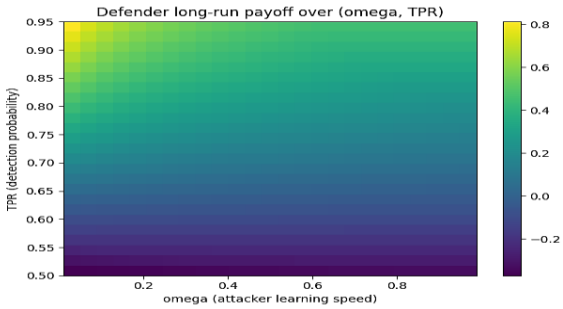


Fig. 6. Heat map of the IDS's long-term payoff as a function of ω and the true positive rate.

Fig. 6 also reveals that being in the secure regime (per *Fig. 5*) does not always mean the IDS's net payoff is maximal. For some (ω, TPR) combinations, stable deterrence is achieved and the attacker's payoff drops sharply, yet the costs of active defense and adaptive learning still eat into the IDS's net payoff, a reminder that "secure" and "economically optimal" are not the same thing. That said, the regions of high IDS payoff in *Fig. 6* line up structurally with the secure regime in *Fig. 5*, which supports the regime analysis: the IDS's payoff peaks mostly where detection quality is high, and the attacker's adaptation speed is kept in check. Payoff maps, then, work well as a quantitative complement to the qualitative regime picture. In short, designing an adaptive IDS has to look past stable security alone and weigh cost–benefit trade-offs at the same time. Pairing regime analysis with long-term payoff maps gives engineers a fuller basis for decision-making, one that can guide practical IDS configurations balanced between security and economic efficiency.

6.5 | Comparison with Existing Approaches

We compared the proposed framework against the classical approaches reported in [6] and [14]; *Table 3* summarizes the comparison.

Table 3. Comparison of the features of the proposed method with previous studies.

Feature	Proposed Method	Method [6]	Method [14]
Modeling detection uncertainty	✓	✗	✗
Adaptive attacker with tunable learning speed	✓	✗	◐
Regime analysis	✓	✗	✗
Memory-one strategy	✓	✗	✓
Bidirectional gradual learning	✓	✗	✗
Modeling false-alarm cost	✓	✗	✗

As *Table 3* shows, the traditional static-game models in [6] and [14] treat the attacker as passive and predictable, while the framework proposed here captures the attacker's dynamic, adaptive behavior directly. Conventional approaches also tend to be overly optimistic about the security state; by explicitly accounting for the attacker's relative learning speed ω, the proposed model draws more realistic security boundaries.

7 | Conclusion and Future Research Directions

This study proposed a repeated-game framework for analyzing how an intrusion detection system and a learning attacker interact under imperfect detection. Its main novelty is modeling the adaptive behavior of both sides at once while folding detection quality directly into the payoff structure. The results confirm that the learning mechanism is numerically stable and that the long-term analysis holds up methodologically. Detection quality, particularly TPR, behaves as a regime-transition parameter, moving the system from attacker dominance through adaptive competition to stable deterrence. Deterrence, however, does not guarantee economic desirability: the cost of active defense can still erode net payoff. And the faster an attacker adapts, the higher the detection quality needed to hold deterrence, which shows that security is a relative notion shaped by the dynamics of the threat itself.

Promising directions for future work include extending the model to multi-attack scenarios, modeling false alarms and delays more precisely, examining convergence indicators in more depth, generalizing to multi-agent settings, and validating the framework in more realistic scenarios. Overall, combining game theory with adaptive learning offers a quantitative way to derive design thresholds for IDS, and shows that stable security calls for attention to detection quality, defense costs, and attacker adaptability all at once.

Statements and Declarations

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

Conceptualization, M.M. and H.S.; Methodology, M.M. and A.K.; Software, M.M.; Validation, M.M., A.K., and H.S.; Formal analysis, M.M.; Investigation, M.M. and A.K.; Writing – original draft, M.M.; Writing – review and editing, A.K. and H.S.; Supervision, H.S. All authors have read and agreed to the published version of the manuscript.

References

[1] Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. *Applied sciences*, *9*(20), 4396. https://doi.org/10.3390/app9204396

[2] Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of information security and applications*, *50*, 102419. https://doi.org/10.1016/j.jisa.2019.102419

[3] Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, *2*(1), 1–22. https://doi.org/10.1186/s42400-019-0038-7%0A%0A

[4] Ayub, M. A., Johnson, W. A., Talbert, D. A., & Siraj, A. (2020, March). Model evasion attack on intrusion detection systems using adversarial machine learning. I*n 2020 54th annual conference on information sciences and systems (CISS) (pp. 1-6). IEEE.* https://doi.org/10.1109/CISS48834.2020.1570617116

[5] Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of network and computer applications*, *60*, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

[6] Manshaei, M. H., Zhu, Q., Alpcan, T., Bacşar, T., & Hubaux, J. P. (2013). Game theory meets network security and privacy. *Acm computing surveys (CSUR)*, *45*(3), 1-39. https://doi.org/10.1145/2480741.2480742

[7] Zhang, H., Mi, Y., Liu, X., Zhang, Y., Wang, J., & Tan, J. (2023). A differential game approach for real-time security defense decision in scale-free networks. *Computer networks*, *224*, 109635. https://doi.org/10.1016/j.comnet.2023.109635

[8] Dehghan, H., & Bigdeli, H. (2026). A Stackelberg signaling game theory model for modeling the defender and attackers under imperfect information situation. *OPSEARCH*, 1–17. https://doi.org/10.1007/s12597-026-01152-8%0A%0A

[9] Chen, J., Xu, Z., Li, Y., Yu, C., Song, J., Yang, H., ... & Wu, Y. (2024, March). Accelerate multi-agent reinforcement learning in zero-sum games with subgame curriculum learning. In *proceedings of the AAAI conference on artificial intelligence* (Vol. 38, No. 10, pp. 11320-11328). https://doi.org/10.1609/aaai.v38i10.29011

[10] Lan, Y., & Hu, Q. (2023). Intrusion detection and response of network node attacks based on game theory. In *proceedings of the 2023 international conference on artificial intelligence, systems and network security* (pp. 386-390). https://doi.org/10.1145/3661638.3661711

[11] Estévez-Pereira, J. J., Fernández, D., & Novoa, F. J. (2020). Network anomaly detection using machine learning techniques. In *Proceedings* (Vol. 54, No. 1, p. 8). MDPI. https://doi.org/10.3390/proceedings2020054008

[12] Apruzzese, G., Andreolini, M., Ferretti, L., Marchetti, M., & Colajanni, M. (2022). Modeling realistic adversarial attacks against network intrusion detection systems. *Digital threats: research and practice (DTRAP)*, *3*(3), 1–19. https://doi.org/10.1145/3469659

[13] Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on emerging telecommunications technologies*, *32*(1), e4150. https://doi.org/10.1002/ett.4150

[14] Ho, E., Rajagopalan, A., Skvortsov, A., Arulampalam, S., & Piraveenan, M. (2022). Game theory in defence applications: A review. *Sensors*, *22*(3), 1032. https://doi.org/10.3390/s22031032

[15] Wang, Z., Xu, S., Xu, G., Yin, Y., Zhang, M., & Sun, D. (2020). Game Theoretical Method for Anomaly-Based Intrusion Detection. *Security and communication networks*, *2020*(1), 8824163. https://doi.org/10.1155/2020/8824163

[16] FathimaAH, N., Khraisat, A., Ibrahim, S., & Li, G. (2025). Adaptive memory replay for network intrusion detection: Tackling data drift and catastrophic forgetting. *Computer networks*, 111712. https://doi.org/10.1016/j.comnet.2025.111712

[17] Malloy, T., & Gonzalez, C. (2023). Learning to defend by attacking (and vice-versa): Transfer of learning in cybersecurity games. In *2023 IEEE european symposium on security and privacy workshops (EuroS&PW)* (pp. 458-464). IEEE. https://doi.org/10.1109/EuroSPW59978.2023.00056