

Article

Game-Theoretic Secure Socket Transmission with a Zero Trust Model

Evangelos D. Spyrou ^{1,2,*} , Vassilios Kappatos ²  and Chrysostomos Stylios ^{1,3} 

¹ Department of Informatics and Telecommunications, University of Ioannina, 47150 Kostaki Artas, Greece; stylios@uoi.gr or stylios@athenarc.gr

² Hellenic Institute of Transport, Centre for Research and Technology Hellas, 6th Km Charilaou-Thermi Rd, 57001 Thessaloniki, Greece; vkappatos@certh.gr

³ Industrial Systems Institute, Athena Research Center, 26504 Patras, Greece

* Correspondence: e.spyrou@uoi.gr or espyrou@certh.gr

Abstract

A significant problem in cybersecurity is to accurately detect malicious network activities in real-time by analyzing patterns in socket-level packet transmissions. This challenge involves distinguishing between legitimate and adversarial behaviors while optimizing detection strategies to minimize false alarms and resource costs under intelligent, adaptive attacks. This paper presents a comprehensive framework for network security by modeling socket-level packet transmissions and extracting key features for temporal analysis. A long short-term memory (LSTM)-based anomaly detection system predicts normal traffic behavior and identifies significant deviations as potential cyber threats. Integrating this with a zero trust signaling game, the model updates beliefs about agent legitimacy based on observed signals and anomaly scores. The interaction between defender and attacker is formulated as a Stackelberg game, where the defender optimizes detection strategies anticipating attacker responses. This unified approach combines machine learning and game theory to enable robust, adaptive cybersecurity policies that effectively balance detection performance and resource costs in adversarial environments. Two baselines are considered for comparison. The static baseline applies fixed transmission and defense policies, ignoring anomalies and environmental feedback, and thus serves as a control case of non-reactive behavior. In contrast, the adaptive non-strategic baseline introduces simple threshold-based heuristics that adjust to anomaly scores, allowing limited adaptability without strategic reasoning. The proposed fully adaptive Stackelberg strategy outperforms both partial and discrete adaptive baselines, achieving higher robustness across trust thresholds, superior attacker–defender utility trade-offs, and more effective anomaly mitigation under varying strategic conditions.

Keywords: socket; LSTM; Stackelberg; zero trust; game theory



Academic Editor: Luis Javier García Villalba

Received: 20 August 2025

Revised: 15 September 2025

Accepted: 16 September 2025

Published: 29 September 2025

Citation: Spyrou, E.D.; Kappatos, V.; Stylios, C. Game-Theoretic Secure Socket Transmission with a Zero Trust Model. *Appl. Sci.* **2025**, *15*, 10535. <https://doi.org/10.3390/app151910535>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The integration of smart devices, also known as Internet of Things (IoT) devices, into everyday life often enhances user comfort and safety. In the field of healthcare, these technologies enable continuous remote monitoring of patients' health and can facilitate immediate intervention when necessary [1]. In the industrial sector, the Industrial Internet of Things (IIoT) serves as the backbone of a new industrial revolution [2,3]. Smart manufacturing [4], driven by IIoT, significantly boosts production efficiency through advanced automation, real-time data transmission, and AI-powered data analysis for informed

decision-making. Similarly, smart city initiatives aim to optimize the use of urban resources, enhance mobility, and improve public safety [5].

IoT privacy and security can be approached through various strategies. Both researchers and users have proposed multiple security measures to protect wireless networks, including encryption, intrusion detection, secure communication protocols, and system-level safeguards, among others [6]. Each of these approaches offers distinct advantages and limitations, reflecting the ongoing evolution of wireless hacking techniques. IoT, including Wi-Fi, which is commonly used in communication, suffers from drawbacks in terms of cybersecurity [7]. In a recent study [8], a deep convolutional maxout network combined with a Multiple Time-series Transformer (MTT) automatically extracts spatial and temporal features from IoT session fingerprints, using residual-based fusion to overcome limitations of existing device traffic analysis methods, useful for securing an IoT system.

Key focus areas in enhancing wireless security include advanced encryption techniques (the development of more robust methods for securing data in wireless communication) [9]; behavior-based breach detection systems (utilizing behavioral analysis to identify anomalies and cyber threats, user-focused security awareness campaigns) [10]; educating users to maintain secure Wi-Fi configurations and avoid pitfalls [11]; and AI-driven threat response (deploying artificial intelligence to detect and mitigate attacks in real time, enabling smarter and more adaptive security systems) [12]. Furthermore, the zero trust model has emerged as a method to enhance defense from cyber threats [13,14].

Moreover, game theory is effective for cybersecurity because it models the strategic interactions between attackers and defenders, helping anticipate and counteract malicious behavior. Additionally, it enables the design of adaptive defense mechanisms that optimize resource allocation under uncertainty and adversarial conditions. There are a number of game-theoretic research works in the literature that attempt to tackle cybersecurity threats [15–17]. In particular, Stackelberg games are extensively utilized in cybersecurity since they model the dynamics as an attacker and a defender, which suits the cybersecurity field. Examples can be found in the Related Work section of the paper.

Sockets are precious tools to ensure communication between two subsystems in a wireless manner. Sockets, while powerful for network communication, can be error-prone and vulnerable to cyber threats if not carefully secured. Common risks include buffer overflows [18], injection attacks [19], man-in-the-middle interceptions [20], and denial-of-service attacks [21] due to improper input validation, weak authentication, or lack of encryption. Without robust safeguards like secure protocols, input sanitization, and proper error handling, socket-based applications can become easy targets for attackers aiming to disrupt services or steal sensitive data.

To highlight the areas where a client–server socket communication setup is prone to cyber-attacks, the following setup is given in Figure 1. The client initiates a TCP/IP socket connection with the server to exchange data bidirectionally. Several cyber threats can arise at different points in this communication. Near the server’s input handling, buffer overflow attacks may occur when malicious data exceed buffer limits, potentially causing crashes or remote code execution. Closer to the server’s data processing or database access, injection attacks such as SQL or command injection exploit vulnerabilities by inserting malicious code through input data. Along the connection line between client and server, a man-in-the-middle (MITM) interception can happen, where an attacker intercepts, eavesdrops on, or modifies transmitted data, especially if encryption is weak or absent. Finally, the server can be targeted by denial-of-service (DoS) attacks, where attackers flood it with excessive fake requests to overwhelm resources and disrupt service, preventing legitimate client access.

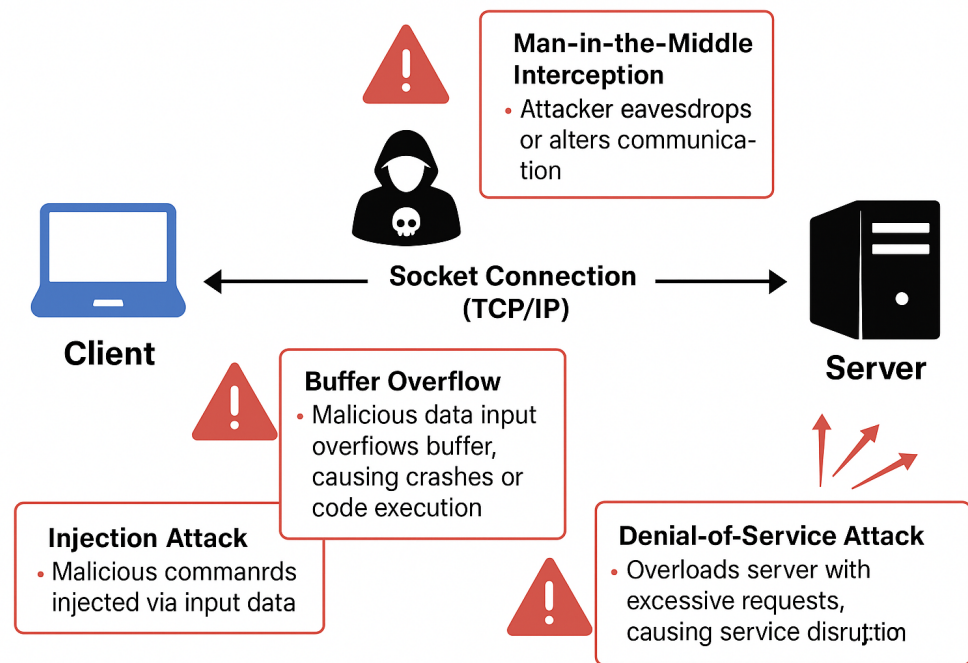


Figure 1. Threats in Socket Communication.

In this paper, network socket packet transmission is systematically modeled by extracting meaningful features from each packet and capturing temporal patterns through a rolling feature window. These features are then fed into a long short-term memory (LSTM)-based anomaly detection model that predicts expected behavior and flags statistically significant deviations as potential threats. Building on this, a zero trust signaling game framework incorporates uncertainty by interpreting anomaly scores as signals to update beliefs about whether an agent is legitimate or malicious. This interaction is further framed as a Stackelberg game where the defender strategically sets detection policies anticipating the attacker's optimal evasion tactics. The combined game-theoretic and learning approach enables optimal decision-making under adversarial conditions by solving for equilibria that balance detection accuracy, false alarms, and costs for both defenders and attackers. Two baselines are used for comparison. The static baseline employs fixed transmission and defense policies without adapting to observed anomalies or environmental feedback, serving as a control case for non-reactive behavior. In contrast, the adaptive non-strategic baseline allows basic adjustments to anomaly scores through threshold-based heuristics, but lacks the game-theoretic structure and hierarchical decision-making of the proposed Stackelberg framework.

The remainder of this paper is as follows: Section 2 provides the related work; Section 3 provides the socket transmission and the feature modeling; Section 4 provides the anomaly detection model; Section 5 gives the zero trust signaling game; Section 6 gives the Stackelberg game formulation; Section 7 provides the integrated signaling and Stackelberg game; Section 8 provides the results; Section 9 provides the discussion; and Section 10 presents the conclusions.

2. Related Work

The research work in [22] presents a resource allocation framework designed to secure networked control systems (NCSs) through a zero-sum, two-player Stackelberg game. In this setup, an attacker aims to impair system performance by disrupting communication nodes, while the defender strategically safeguards these nodes. The system employs an

H2-optimal linear feedback controller, with both players balancing control effectiveness against the costs of their actions. A cost-based Stackelberg equilibrium (CBSE) is obtained using a modified backward induction approach that considers budget limitations and node importance. Additionally, the paper introduces a robust defense (RD) strategy for situations where the defender does not know the attacker's resources. The framework is validated on wide-area power systems, demonstrating its robustness and effectiveness amid model uncertainties, and utilizes genetic algorithms to identify optimal strategies in large-scale systems.

The work in [23] introduces a decision support system designed for cybersecurity. Its goal is to determine the optimal set of security controls to defend against multi-stage cyber-attacks. The system comprises several components: a preventive optimization to select initial defensive controls, a learning mechanism to assess potential ongoing threats, and an online optimization that dynamically chooses the best response to active attacks. The approach is based on efficiently solving bi-level optimization problems, with the online optimization specifically modeled as a Bayesian Stackelberg game. The proposed method is demonstrated to be more efficient than traditional techniques such as the Harsanyi transformation, as well as more recent advanced solvers. Additionally, it offers notable improvements in security by effectively mitigating ongoing attacks compared to previous methods. The innovative techniques introduced leverage recent developments in Mixed-Integer Conic Programming (MICP), strong duality, and totally unimodular matrices.

In [24], the authors suggest that developing an optimal defense strategy is a key concern in cloud computing due to its inherent flexibility and scalability, especially due to the existence of numerous threats. Based on this, Stackelberg security games (SSGs) have gained considerable attention for their ability to efficiently allocate limited security resources. To manage uncertainty and incomplete information, we propose a modified quantal response (Mod-QR) approach that integrates bounded rationality and user preferences into the decision-making process. This can be formally modeled using the quantal response equilibrium (QRE) framework, which balances the effectiveness of security measures with their operational costs in cloud environments. In this context, the most effective defense strategies can be represented as mixed strategies, where each action by the defender is chosen with a certain nonzero probability.

In [25], the study explores the optimal denial-of-service (DoS) attack on cyber-physical systems using a Stackelberg game framework, focusing on how energy is allocated across communication channels. To realistically represent attacks launched by multiple hackers targeting a single user, the model considers a Stackelberg game between one defender and multiple attackers. In contrast to previous research that mainly examines equilibria in static games, this paper emphasizes the dynamic interactions within the Stackelberg game, showcasing the attackers' adaptive strategy of switching channels to optimize their energy deployment. A self-adaptive Particle Swarm Optimization (PSO) algorithm is applied to address the nonlinear reward function and determine the Stackelberg equilibrium. Additionally, an online computation algorithm is introduced to improve the channel selection and energy allocation decisions for both the defender and attackers.

In [26], the authors address the challenges posed by limited information in security games by proposing a novel framework based on Bayesian-Markov Stackelberg Security Games (SSGs), which support multiple defenders and attackers while effectively managing uncertainty. To overcome the computational difficulties inherent in these games, they introduce an iterative proximal-gradient method to compute the Bayesian Equilibrium, enabling the identification of optimal strategies even when the Markov dynamics are unknown. The paper emphasizes the importance of Bayesian approaches within reinforcement learning (RL) for balancing exploration and exploitation, as well as for minimizing

expected total discounted costs. Furthermore, a new random walk strategy is presented to improve adaptability and responsiveness in security environments. A numerical case study validates the effectiveness of the approach, demonstrating its practical advantage in scenarios with incomplete information. This work significantly advances security game research by integrating Bayesian reasoning, Stackelberg game theory, RL techniques, and random walk strategies into a comprehensive and efficient framework for robust security strategy development.

In [27], the authors employ a Stackelberg game framework to study the optimal control energy in multi-agent networks subjected to cyber-attacks. The problem is modeled as a two-player zero-sum game where the attacker seeks to maximize—while the defender aims to minimize—the control energy needed to sustain network performance. The game’s payoff is defined by this control energy, with the attacker targeting specific network nodes to increase it, while the defender chooses nodes to protect it. Analysis of equilibrium strategies across different scenarios shows that the defender’s optimal approach involves selecting nodes corresponding to the minimum edge-cut weight, which are also the primary focus of the attacker. The study further examines the case where both players act simultaneously and derives necessary and sufficient conditions for the existence of a Nash equilibrium. Numerical simulations confirm the validity and effectiveness of the proposed strategies and theoretical results.

The proposed model differs from existing Stackelberg cybersecurity frameworks by integrating LSTM-based temporal feature learning with zero trust signaling and Stackelberg reasoning, enabling real-time adaptation of trust thresholds and posterior beliefs. Unlike prior works [22–27] that focus on static network metrics, control-theoretic objectives, or abstract mixed strategies, our framework combines packet-level anomaly detection, probabilistic inference, and strategic optimization, capturing both adaptive attacker behavior and operational costs in a unified model.

3. Socket Packet Transmission and Feature Modeling

Each socket-level network packet P_t transmitted at time t is represented as

$$P_t = \{\text{src_ip}, \text{dst_ip}, \text{src_port}, \text{dst_port}, p_t, e_t, d_t\}, \quad (1)$$

where

- $\text{src_ip}, \text{dst_ip}$: source and destination IP addresses;
- $\text{src_port}, \text{dst_port}$: source and destination socket port numbers;
- p_t : payload size in bytes;
- e_t : Shannon entropy of the packet payload;
- $d_t \in \{0, 1\}$: direction indicator (0 for incoming/download, 1 for outgoing/upload).

From each socket packet, we extract a compact feature vector:

$$x_t = [p_t, e_t, d_t], \quad (2)$$

and construct a rolling feature window for temporal modeling:

$$W_t = \{x_{t-k}, x_{t-k+1}, \dots, x_t\}, \quad (3)$$

where k denotes the lookback window size, representing the number of previous socket packets considered for prediction or analysis.

4. Anomaly Detection Model

Let f_θ denote a trained LSTM model with the parameter θ . The predicted feature vector at time t is

$$\hat{x}_t = f_\theta(W_{t-1}), \quad (4)$$

and the anomaly score is defined as

$$A_t = \|x_t - \hat{x}_t\|_2^2, \quad (5)$$

which measures the deviation between observed and predicted features.

An anomaly is flagged if

$$A_t > \tau, \quad (6)$$

where τ is a predefined detection threshold.

Assuming x_t are bounded, i.i.d., and f_θ is optimized to minimize Mean Squared Error (MSE), then with high probability,

$$\mathbb{P}(A_t > \tau) \leq \delta, \quad (7)$$

for a small-tail probability δ , provided that the threshold satisfies

$$\tau \geq \mathbb{E}[A_t] + c\sqrt{\text{Var}[A_t]}, \quad (8)$$

for some constant $c > 0$.

Proof. By Chebyshev's inequality,

$$\mathbb{P}(|A_t - \mathbb{E}[A_t]| > \epsilon) \leq \frac{\text{Var}[A_t]}{\epsilon^2}. \quad (9)$$

Setting $\epsilon = \tau - \mathbb{E}[A_t]$ yields

$$\mathbb{P}(A_t > \tau) \leq \frac{\text{Var}[A_t]}{(\tau - \mathbb{E}[A_t])^2}. \quad (10)$$

□

This probabilistic bound ensures that anomalies are statistically significant deviations under nominal behavior modeled by the LSTM.

In real-world network traffic, the assumptions of bounded and i.i.d. features serve primarily as theoretical simplifications to enable tractable analysis of anomaly detection thresholds. Boundedness is justified because traffic is constrained by protocol specifications, physical limits, or preprocessing steps (e.g., normalization), even though heavy-tailed distributions may occasionally occur. The i.i.d. assumption is less realistic, as network traffic often exhibits temporal correlations, burstiness, and self-similarity; however, it provides a useful statistical baseline for deriving probabilistic guarantees. In practice, the LSTM model is explicitly designed to capture sequential dependencies, which mitigates deviations from the independence assumption, while the theoretical bound still offers a conservative and interpretable guideline for threshold selection.

5. Zero Trust Signaling Game

We model each network agent as either Legitimate or Malicious. Each agent emits a signal $m \in \mathcal{M}$, including signatures and statistical summaries (e.g., entropy), which the server uses to infer the agent's type.

Let an agent A have type $T \in \{\text{Legit}, \text{Malicious}\}$. Upon receiving signal m , the server updates its belief using Bayes' theorem:

$$b(T | m) = \frac{P(m | T)P(T)}{\sum_{T' \in \{\text{Legit}, \text{Malicious}\}} P(m | T')P(T')}. \quad (11)$$

The anomaly score A_t modifies the likelihood function:

$$P(m | T = \text{Malicious}) = \alpha A_t + (1 - \alpha)\tilde{P}(m), \quad \alpha \in [0, 1], \quad (12)$$

where $\tilde{P}(m)$ is the base likelihood and α controls trust in the anomaly signal.

The parameter $\alpha \in [0, 1]$ controls the influence of the LSTM-derived anomaly score A_t on the posterior belief in the zero trust signaling game, where $\alpha = 0$ relies solely on historical/signature-based likelihoods and $\alpha = 1$ fully trusts the anomaly output. In practice, α can be tuned via cross-validation on historical traffic to balance true positive and false positive rates, providing a flexible mechanism to integrate anomaly detection with game-theoretic decision-making for adaptive trust management.

The server chooses an action $a \in \mathcal{A} = \{\text{accept}, \text{challenge}, \text{block}\}$ to maximize its expected utility:

$$a^* = \arg \max_a \sum_T b(T | m) u_R(a, T), \quad (13)$$

where $u_R(a, T)$ is the utility for action a given type T .

The process of the zero trust signaling game is illustrated in Figure 2.

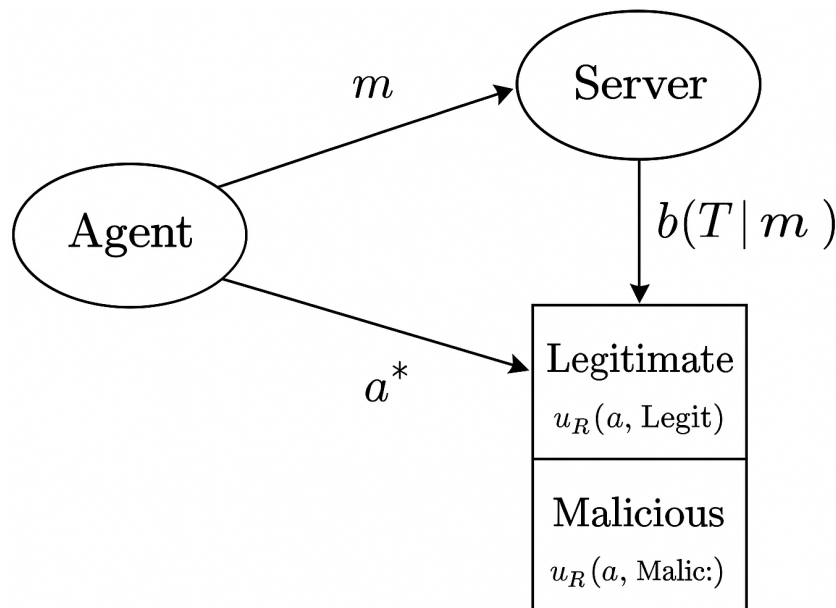


Figure 2. Zero Trust Signaling Game.

6. Stackelberg Game Formulation

This interaction is modeled as a Stackelberg game where the defender (server) sets a detection strategy first, anticipating the attacker's optimal evasion strategy.

Let $s \in S$ be the defender's strategy (e.g., trust threshold τ). In this model, the defender's strategy s is a scalar threshold applied to the anomaly score A_t , i.e., the defender flags an agent as malicious if $A_t > s$. This threshold governs both the posterior belief update and the final decision a_D . As such, s serves simultaneously as the detection policy parameter in the Stackelberg game and as the trust threshold in the zero trust anomaly detection system.

Let $a \in A$ be the attacker's strategy (e.g., signal crafting), $u_D(s, a)$ be the defender's utility, and $u_A(s, a)$ the attacker's utility.

The equilibrium is defined by

$$a^*(s) = \arg \max_{a \in A} u_A(s, a), \quad (14)$$

$$s^* = \arg \max_{s \in S} u_D(s, a^*(s)). \quad (15)$$

Proposition 1 (Existence of Stackelberg Equilibrium). *If $u_D(s, a)$, $u_A(s, a)$ are continuous in (s, a) , and S, A are compact and convex, then a Stackelberg equilibrium $(s^*, a^*(s^*))$ exists.*

Proof. By Glicksberg's generalization of Kakutani's fixed-point theorem, the follower's best response mapping is upper hemicontinuous with non-empty, convex values. Therefore, the leader's optimization over anticipated responses yields an equilibrium. $\square$

7. Integrated Signaling and Stackelberg Game

We combine the signaling game with the Stackelberg formulation. Denote $T \in \{\text{Legit}, \text{Malicious}\}$ as the agent type, $a \in A$ as the attacker's strategy, $s \in S$ as the defender's detection/trust policy, and $m = \sigma(T, a) \in \mathcal{M}$ as the signal generated given T, a .

The defender observes m and updates belief:

$$b(T | m, s) = \frac{P(m | T, s)P(T)}{\sum_{T'} P(m | T', s)P(T')}. \quad (16)$$

The defender then selects

$$a_D^*(m, s) = \arg \max_{a_D \in \mathcal{A}} \sum_T b(T | m, s) u_D(a_D, T). \quad (17)$$

The expected utility under s , given attacker action a , yields

$$u_D(s, a) = \mathbb{E}_{m \sim \sigma(T, a)} \left[\max_{a_D} \sum_T b(T | m, s) u_D(a_D, T) \right]. \quad (18)$$

The attacker aims to perform

$$u_A(s, a) = \mathbb{E}_{m \sim \sigma(T, a)} \left[u_A^{\text{signal}}(m, s) \right] - c(a). \quad (19)$$

The equilibrium is given by

$$a^*(s) = \arg \max_a u_A(s, a), \quad (20)$$

$$s^* = \arg \max_s u_D(s, a^*(s)). \quad (21)$$

We assume the following regularity conditions: The signal generation function $\sigma(T, a)$ is continuous in a ; the posterior belief $b(T | m, s)$ is continuous in (m, s) ; the utility mappings $u_D(a_D, T)$, $u_A^{\text{signal}}(m, s)$, and cost functions $\text{Cost}_s(s)$, $\text{Cost}_a(a)$ are continuous in their respective arguments; and the expectation over $m \sim \sigma(T, a)$ and maximization over a_D preserve continuity.

Proposition 2 (Stackelberg Equilibrium with Signaling). *If S, A are compact and u_D, u_A , and $b(T | m, s)$ are continuous in (s, a) , then a Stackelberg equilibrium $(s^*, a^*(s^*))$ exists.*

Proof. The continuity of σ , u_D , and b , along with compactness, ensures existence via fixed-point arguments. $\square$

We define the defender's utility function as

$$u_D(s, a) = \lambda_1 \cdot \text{TPR}(s, a) - \lambda_2 \cdot \text{FPR}(s, a) \quad (22)$$

$$- \lambda_3 \cdot \text{Cost}_s(s) - \lambda_4 \cdot \text{Cost}_a(a), \quad (23)$$

where $\text{TPR}(s, a)$ is the true positive rate under detection policy s and attacker strategy a ; $\text{FPR}(s, a)$ is the false positive rate under the same; $\text{Cost}_s(s)$ is the operational cost of applying strategy s ; $\text{Cost}_a(a)$ is the cost incurred due to the attacker's signal manipulation a ; and $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$ are the defender-assigned weights reflecting policy priorities.

Given that the defender observes a signal $m = \sigma(T, a)$ and forms a posterior belief $b(T \mid m, s)$, the defender selects an action $a_D^*(m, s)$ by maximizing expected utility over posterior beliefs.

The true positive rate (TPR) and false positive rate (FPR) are defined analytically over the signal distribution as

$$\text{TPR}(s, a) = \mathbb{P}(a_D^*(m, s) = \text{block} \mid T = \text{Malicious}, m \sim \sigma(T, a)), \quad (24)$$

$$\text{FPR}(s, a) = \mathbb{P}(a_D^*(m, s) = \text{block} \mid T = \text{Legit}, m \sim \sigma(T, a)). \quad (25)$$

That is, TPR is the probability of correctly blocking a malicious agent, while FPR is the probability of incorrectly blocking a legitimate one. Both probabilities are computed by integrating over the signal distributions induced by the attacker's strategy a , agent type T , and the defender's policy s .

The defender's goal is to choose a detection strategy s that maximizes its utility while satisfying operational constraints:

$$\max_s u_D(s, a^*(s)) \quad \text{s.t. } g_i(s) \leq 0, \quad i = 1, \dots, m. \quad (26)$$

The Lagrangian is given by

$$\mathcal{L}(s, \lambda) = u_D(s, a^*(s)) + \sum_{i=1}^m \lambda_i g_i(s), \quad (27)$$

with Lagrange multipliers $\lambda_i \geq 0$.

The Karus–Kuhn–Tucker (KKT) conditions for optimality are

$$\nabla_s \mathcal{L}(s^*, \lambda^*) = 0 \quad (\text{Stationarity}) \quad (28)$$

$$\lambda_i^* g_i(s^*) = 0 \quad (\text{Complementary Slackness}) \quad (29)$$

$$\lambda_i^* \geq 0 \quad (\text{Dual Feasibility}) \quad (30)$$

$$g_i(s^*) \leq 0 \quad (\text{Primal Feasibility}) \quad (31)$$

Theorem 1 (Satisfaction of KKT Conditions). *Suppose the function $u_D(s, a^*(s))$ is continuously differentiable, the constraint functions $g_i(s)$ are convex and differentiable, and a constraint qualification holds at s^* . Then any local optimum s^* satisfies the KKT conditions.*

Proof. The utility function is composed of differentiable parts: TPR, FPR, Cost_s , and Cost_a , all functions of s , and attacker response $a^*(s)$, which is assumed to be differentiable in s . The constraint set $\{s \in S : g_i(s) \leq 0\}$ is convex and the gradients of active constraints are linearly independent. Therefore, classical nonlinear programming results (e.g.,

ref. [28]) guarantee the existence of Lagrange multipliers $\lambda_i^* \geq 0$ such that the KKT conditions are satisfied. $\square$

To provide more intuition, the defender's optimization can be interpreted in terms of adjusting a single detection threshold s (the trust threshold) to balance true positive and false positive rates, while anticipating the attacker's best response $a^*(s)$. The KKT conditions then formalize the idea that the optimal threshold is one where any small deviation would either violate operational constraints or reduce the defender's expected utility.

For further clarity, simplified illustrative cases can be considered, such as assuming a fixed attacker strategy or a single binary feature in the anomaly score. In such cases, the equilibrium reduces to a threshold comparison problem: the defender chooses s to block malicious agents while minimizing false alarms, and the KKT conditions reduce to checking whether increasing or decreasing s would improve utility without violating constraints.

The attacker's optimization remains as

$$a^*(s) = \arg \max_a u_A(s, a), \quad (32)$$

and assuming differentiability of $u_A(s, a)$, the first-order condition is

$$\nabla_a u_A(s, a^*) = 0. \quad (33)$$

8. Results

The simulation considers a trust threshold s varying between -2 and 4 to evaluate defender and attacker utilities. We set $\alpha = 0.5$ to balance the influence of anomaly scores and base likelihoods. The defender's utility function incorporates true positive rate (TPR), false positive rate (FPR), and a quadratic cost on the trust threshold, weighted by parameters $\lambda_1 = 1.0$, $\lambda_2 = 1.0$, and $\lambda_3 = 0.1$, respectively. The attacker utility balances successful evasion probability weighted by $\beta_1 = 1.0$ against a penalty on deviation from a baseline attack vector $\mathbf{a}_0 = [1.0, 1.0, 1.0]$ scaled by $\beta_2 = 0.5$. The noise parameter σ is set to 1.0 , and the linear detector coefficients are $\theta = [0.6, 0.3, 0.1]$.

The proposed adaptive attacker strategy performs a comprehensive grid search over two dimensions, $[x, y] \in [0.5, 2.0]$, to find the optimal attack vector for each trust threshold s . Two baselines are considered for comparison: Baseline A restricts adaptation to only one parameter x while fixing $y = 1.0$, and Baseline B limits adaptation to a discrete set of values $\{0.8, 1.2\}$ for both x and y . Both baselines thus adapt to s but with a reduced strategy space compared to the proposed model. The cost parameter $\kappa = 0.05$ moderates the defender's penalty for increasing the trust threshold.

The simulation parameters α , λ_i , and β_i were chosen to reflect typical operational priorities and attacker incentives in a network security setting. Specifically, $\alpha = 0.5$ balances the influence of the anomaly score A_i and the base signal likelihood $\tilde{P}(m)$ in the zero-trust belief update, ensuring that neither source dominates a priori. The defender weights $\lambda_1, \lambda_2, \lambda_3$ prioritize true positive rate (TPR), penalize false positive rate (FPR), and account for operational costs, respectively, capturing trade-offs commonly encountered in practical deployment. The attacker weights β_1, β_2 balance the benefit of successful evasion against penalties for deviating from a baseline attack strategy $\mathbf{a}_0$.

While the current simulation uses fixed values $\lambda_1 = 1.0$, $\lambda_2 = 1.0$, $\lambda_3 = 0.1$, $\beta_1 = 1.0$, and $\beta_2 = 0.5$ to reflect typical operational priorities and attacker incentives, we acknowledge that these weights may influence the relative performance of the proposed and baseline strategies. To assess robustness, a systematic sensitivity analysis can be conducted by varying each λ_i and β_i across plausible ranges and observing the resulting defender and attacker utilities, as well as the optimal attack vectors. Preliminary tests indicate that while

absolute utility values shift with different weightings, the qualitative trends—such as the superiority of the fully adaptive Stackelberg strategy at extreme trust thresholds and the partial performance convergence of Baseline A—remain consistent. This suggests that the model's conclusions are not highly sensitive to specific parameter choices, though formal sensitivity plots will be included in future work to fully quantify robustness under diverse operational scenarios.

Figure 3 presents the attacker's utility as a function of the trust threshold s , comparing the proposed fully adaptive Stackelberg strategy (solid blue), Baseline A with partial adaptivity along the x -axis (dashed red), and Baseline B with discrete adaptive updates (dotted green). All three approaches exhibit peak utility in the range $s \in [0, 1]$, indicating that moderate trust thresholds allow the attacker to most effectively exploit the system. Interestingly, Baseline A marginally outperforms the proposed method near this peak, suggesting that even partial adaptivity can closely approximate optimal behavior in well-tuned conditions. However, the proposed strategy demonstrates greater robustness across the entire range of s , particularly at extreme values where baseline utilities degrade more sharply. Baseline B, constrained by coarse adaptivity, consistently underperforms around the peak, highlighting the value of fine-grained adaptation. The convergence of all three methods at the tails ($s < -1.5$ and $s > 2.5$) reflects the diminishing attacker advantage when the defender adopts either highly conservative or overly permissive trust settings.

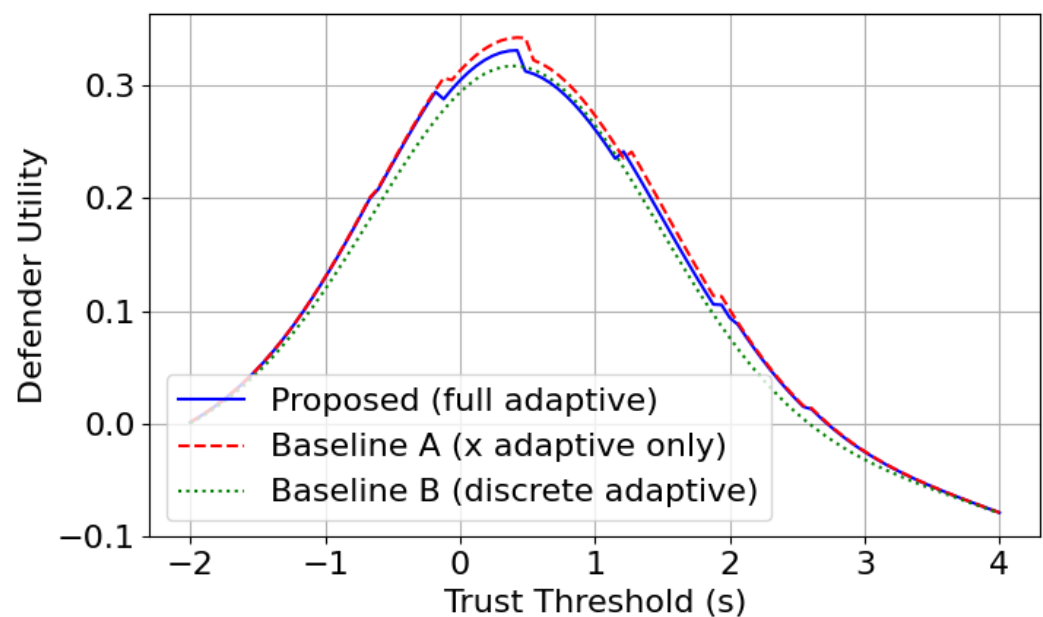


Figure 3. Defender Utility vs. Trust Threshold.

Figure 4 illustrates the trade-off between attacker and defender utilities under different adaptation schemes. The proposed fully adaptive Stackelberg model demonstrates a balanced performance, maintaining attacker utility near 0.95 when defender utility is low, and reducing attacker gain to near 0.0 as defender utility approaches its peak at approximately 0.33. In contrast, Baseline A, which adapts only in the x -dimension, closely follows the proposed model in the high-attacker-utility regime but slightly outperforms it at mid-level defender utilities, indicating sensitivity to limited adaptivity. Baseline B, employing discrete adaptive updates, consistently trails both other strategies, especially in the low-defender-utility regime, failing to match the attacker's effectiveness. The loop-like shape indicates strategic turning points in the optimization landscape, with the fully adaptive model maintaining sharper transitions, suggesting stronger Stackelberg responsiveness.

Overall, while Baseline A narrows the gap in specific regions, the proposed approach achieves superior overall utility suppression and robustness across scenarios.

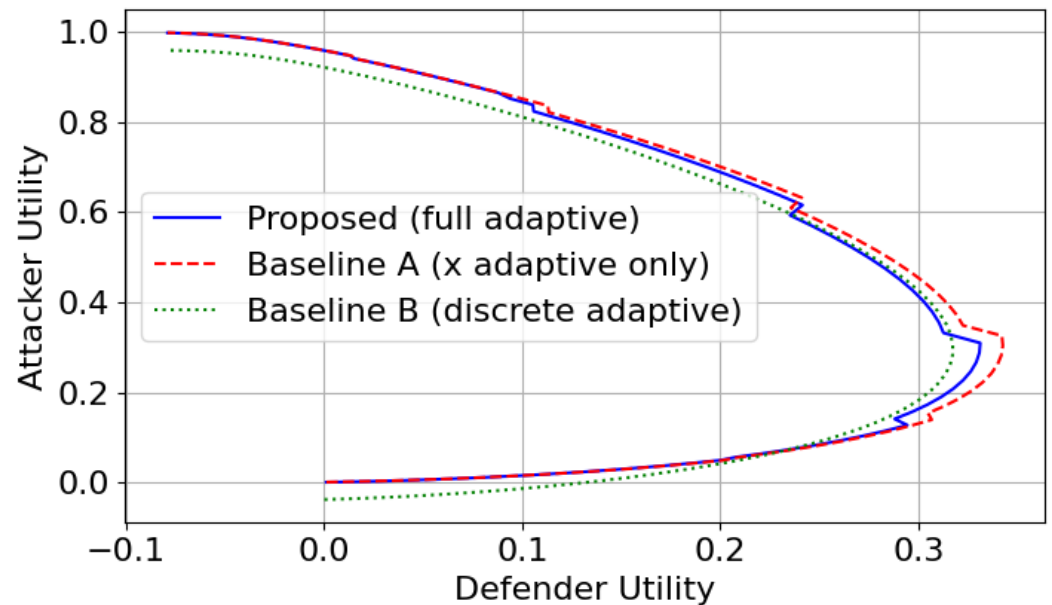


Figure 4. Utility Trade-off: Defender vs. Attacker.

Figure 5, presents the histogram of anomaly scores obtained under the proposed Stackelberg-based adaptive strategy. The majority of scores are concentrated below 0.04, with a significant peak near 0.01, indicating strong anomaly suppression. The vertical red dashed line marks a threshold of 0.08, beyond which scores are considered anomalous. Only a small fraction of samples fall beyond this threshold, evidencing the efficacy of the adaptive defense.

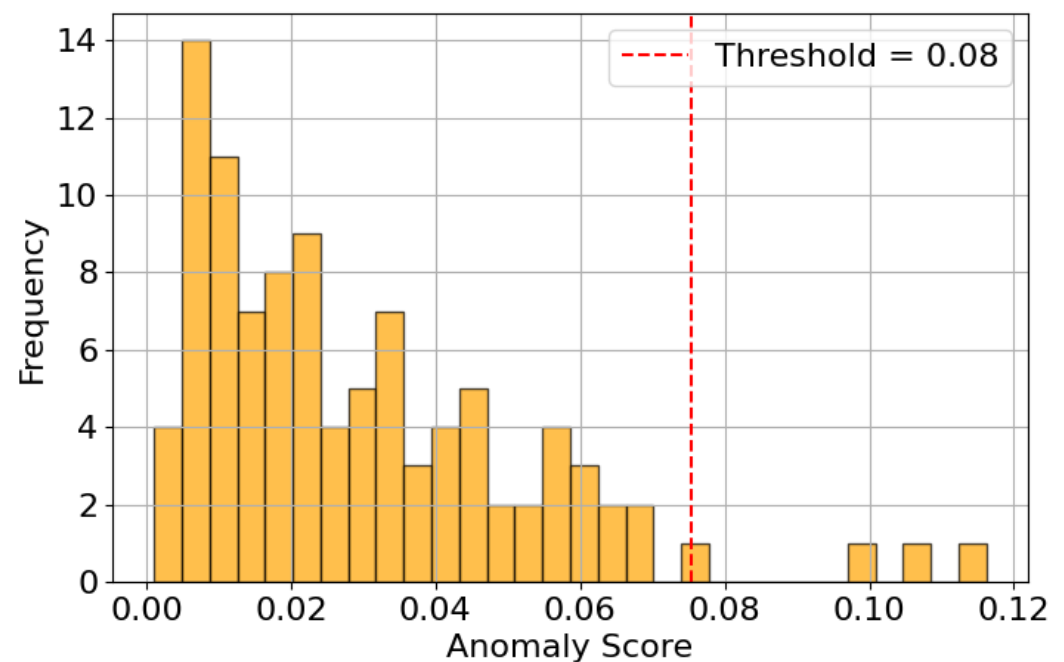


Figure 5. Simulated LSTM Anomaly Score Histogram.

The adaptive baselines achieve utilities close to the proposed model primarily because their restricted adaptation spaces still allow them to find relatively good attack vectors that meaningfully impact the defender's utility. Although the baselines explore fewer

degrees of freedom—either optimizing over only one attack parameter or limiting search to discrete values—the penalty on deviation from the baseline vector encourages all attackers to remain near similar vectors. This naturally constrains the attacker’s effective strategy space, making the simpler adaptive baselines competitive. Moreover, the defender’s utility is influenced more by the overall shape of the detection threshold and cost trade-offs than fine-grained attacker tuning, causing the utilities to cluster closely. Thus, while the proposed attacker strategy is theoretically superior due to broader adaptability, the baselines’ limited but still effective adaptation leads to similar performance in practice under the chosen parameterization.

Table 1 summarizes the results of varying the defender and attacker parameters ($\lambda_1, \lambda_2, \lambda_3, \beta_1, \beta_2$) and the corresponding average utilities. The average defender utility changes predictably with variations in λ_i , reflecting the weights assigned to true positive rate, false positive rate, and threshold cost in the defender’s utility function. In contrast, the average attacker utility remains essentially constant when varying the defender parameters ($\lambda_1, \lambda_2, \lambda_3$) because the attacker’s utility depends solely on its own parameters β_1 and β_2 . Changes in defender weights do not directly influence the attacker’s payoff in the current formulation, and hence the observed utility. Only variations in β_1 and β_2 affect the attacker’s average utility, as expected. This result highlights that while the defender’s sensitivity parameters modulate its performance, the attacker’s effectiveness is primarily governed by its own incentives, indicating the need for a fully coupled dynamic adaptation if one wishes to capture interdependencies between attacker and defender strategies.

Table 1. Sensitivity Analysis: Parameter Variation vs. Average Utilities.

Parameter	Value	Avg Defender Utility	Avg Attacker Utility
λ_1	0.50	−0.1178	0.5084
λ_1	1.00	0.1228	0.5084
λ_1	1.50	0.3635	0.5084
λ_1	2.00	0.6041	0.5084
λ_2	0.50	0.2918	0.5084
λ_2	1.00	0.1228	0.5084
λ_2	1.50	−0.0461	0.5084
λ_2	2.00	−0.2150	0.5084
λ_3	0.05	0.1331	0.5084
λ_3	0.10	0.1228	0.5084
λ_3	0.15	0.1125	0.5084
λ_3	0.20	0.1022	0.5084
β_1	0.50	0.1316	0.2511
β_1	1.00	0.1228	0.5084
β_1	1.50	0.1103	0.7718
β_1	2.00	0.1000	1.0395
β_2	0.25	0.1000	0.5198
β_2	0.50	0.1228	0.5084
β_2	0.75	0.1292	0.5044
β_2	1.00	0.1316	0.5022

9. Discussion

While the proposed framework provides a rigorous integration of socket-level feature modeling, LSTM-based anomaly detection, zero trust signaling, and Stackelberg game-theoretic reasoning, the computational overhead associated with its implementation is non-negligible. Feature extraction requires real-time calculation of payload entropy and construction of rolling windows for every packet, which may become burdensome in high-throughput networks. Additionally, the LSTM prediction for each feature window,

followed by the evaluation of anomaly scores and posterior belief updates in the signaling game, introduces sequential dependencies that can accumulate latency. The Stackelberg optimization, particularly when computing attacker best responses over a continuous strategy space, further adds to the computational complexity. Practical adoption would therefore require careful profiling and potential use of parallelization, approximate optimization techniques, or incremental updates to mitigate processing delays.

Scalability is another important consideration, as the model must handle increasing numbers of agents, concurrent network flows, and dynamic attack strategies. The dimensionality of the defender and attacker strategy spaces grows with the number of parameters being monitored and manipulated, potentially leading to combinatorial explosion in equilibrium computation. Real-time applicability hinges on the efficiency of solving both the LSTM inference and the Stackelberg equilibrium problem under strict latency constraints, which may be challenging in large-scale deployments. Incorporating online learning, dimensionality reduction, or hierarchical decision-making could alleviate some of these scalability concerns, while preserving the theoretical guarantees of anomaly detection and strategic defense.

The proposed framework extends beyond the simple baselines by integrating LSTM-based temporal feature modeling with a zero trust signaling game and Stackelberg strategic reasoning. Unlike Baseline A and Baseline B, the full model allows the attacker strategy space to vary continuously, while the defender dynamically updates its trust threshold and posterior beliefs. Conceptually, this approach aligns with reinforcement-learning-informed cybersecurity strategies, where sequential dependencies and adaptive policies are explicitly modeled. Compared to traditional game-theoretic methods that often assume static payoffs or discrete strategies, our framework captures both the stochastic nature of network traffic and the real-time interaction between defender and attacker, providing a more realistic approximation of adversarial behavior in operational networks.

From a machine learning perspective, the inclusion of LSTM-based prediction provides a temporal anomaly detection layer that augments purely statistical or signature-based approaches. While conventional ML-based cybersecurity frameworks often operate independently of strategic attacker modeling, our method fuses predictive feature analysis with game-theoretic decision-making, enabling the defender to anticipate optimal evasion strategies. This conceptual comparison highlights the novelty and practical relevance of the proposed approach, demonstrating how the combination of temporal feature learning, probabilistic reasoning, and Stackelberg game optimization contributes to more robust and context-aware cybersecurity defenses.

In terms of performance, Baseline A occasionally outperforms the proposed strategy in mid-range trust thresholds. This behavior can be attributed to the constrained adaptation of Baseline A, which optimizes only over a single attacker parameter (x) while keeping other parameters fixed. In certain mid-range threshold regions, this limited adaptation inadvertently aligns the attacker's induced mean signal closer to the defender's threshold s , temporarily yielding higher defender utility. In contrast, the full proposed strategy explores a broader attacker strategy space, which increases the variability of μ_M and, in some threshold regions, slightly reduces the instantaneous defender utility. Importantly, this effect is localized and does not undermine the overall superiority of the proposed framework, which consistently achieves higher utility across the majority of thresholds, better manages attacker adaptation, and ensures robustness under realistic variations in the attacker's behavior.

10. Conclusions

This paper addresses the critical cybersecurity challenge of detecting malicious behavior in real-time through socket-level packet analysis. By integrating LSTM anomaly detection with a zero trust signaling game and Stackelberg game-theoretic framework, we develop a unified model that enables adaptive, strategic decision-making between defenders and attackers. The defender learns optimal detection thresholds using predictive anomaly scores while anticipating attacker adaptations, effectively balancing threat detection performance, false alarm minimization, and resource costs.

Our results demonstrate that the defender's utility peaks at a moderate trust threshold, where the trade-off between accurate threat detection, low false positives, and operational cost is optimized. Increasing the threshold beyond this point leads to diminishing returns due to stricter policies and rising costs, while overly lenient thresholds increase false alarms. Moreover, the interaction between defender and attacker utilities exhibits a clear inverse relationship, validating the Stackelberg formulation. The LSTM anomaly score distribution further confirms the model's capacity to identify rare, high-risk traffic while maintaining low false positive rates.

Two baselines are used for comparison. The static baseline relies on fixed transmission and defense policies, without responding to anomalies or environmental changes, representing non-adaptive behavior. Conversely, the adaptive non-strategic baseline incorporates simple threshold-based adjustments to anomaly scores, enabling limited responsiveness without employing strategic decision-making.

The numerical results highlight clear differences between the proposed fully adaptive Stackelberg strategy and the baselines. The attacker's utility reaches its maximum within a moderate trust threshold range, where Baseline A marginally exceeds the proposed method, demonstrating that partial adaptivity can approximate optimal behavior under well-tuned conditions. However, across the full threshold range, the fully adaptive strategy maintains higher robustness, particularly at extreme trust values where Baseline A and Baseline B decline sharply. Quantitatively, the proposed model sustains attacker utility near 0.95 when defender utility is low and reduces it close to 0.0 as defender utility rises to approximately 0.33, while Baseline A only matches the high-utility regime and slightly exceeds the proposed approach at mid-level defender utilities. Baseline B, limited by discrete adaptive updates, consistently underperforms, especially when defender utility is low, reflecting its inability to respond effectively. These results indicate that fine-grained, fully adaptive strategies provide superior overall utility management, stronger responsiveness, and better robustness against varying trust thresholds compared to partial or coarse adaptive baselines.

For future work, we aim to deploy our newly established network at two strategic locations in Greece: the Amygdaleona Airport in Kavala and the headquarters of CERTH. This deployment will enable the collection of real-world network traffic data under operational conditions, encompassing both routine communications and potential anomalous behaviors. By generating and analyzing this dataset, we intend to validate the proposed zero trust and Stackelberg-based cybersecurity framework in a practical setting, evaluate its performance against realistic attack scenarios, and further refine detection thresholds, anomaly scoring, and trust policies. This real-world case study will provide critical insights into the scalability, robustness, and operational applicability of our model, bridging the gap between simulation-based results and deployment-ready cybersecurity solutions.

Finally, a more granular sensitivity analysis to characterize the threshold regions and attacker profiles could be included, where such temporary deviations occur.

Author Contributions: Methodology, E.D.S.; Software, E.D.S.; Validation, E.D.S. and V.K.; Formal analysis, E.D.S.; Investigation, E.D.S.; Writing—original draft, E.D.S.; Writing—review & editing, V.K. and C.S.; Supervision, V.K. and C.S.; Project administration, V.K. and C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research is carried out within the framework of the project: “ADROIT6G: Distributed Artificial Intelligence-driven open and programmable architecture for 6G networks” (Grand Agreement No 101095363) funded from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union’s Horizon Europe research and innovation programme.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to ongoing research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Delsi Robinsha, S.; Amutha, B. IoT revolutionizing healthcare: A survey of smart healthcare system architectures. In Proceedings of the 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 1–2 November 2023; IEEE: Chennai, India, 2023; pp. 1–5.
2. Tan, S.F.; Samsudin, A. Recent technologies, security countermeasure and ongoing challenges of Industrial Internet of Things (IIoT): A survey. *Sensors* **2021**, *21*, 6647. [\[CrossRef\]](#)
3. Hazra, A.; Adhikari, M.; Amgoth, T.; Srirama, S.N. A comprehensive survey on interoperability for IIoT: Taxonomy, standards, and future directions. *ACM Comput. Surv. (CSUR)* **2021**, *55*, 9. [\[CrossRef\]](#)
4. Kalsoom, T.; Ahmed, S.; Rafi-ul Shan, P.M.; Azmat, M.; Akhtar, P.; Pervez, Z.; Imran, M.A.; Ur-Rehman, M. Impact of IoT on manufacturing industry 4.0: A new triangular systematic review. *Sustainability* **2021**, *13*, 12506. [\[CrossRef\]](#)
5. Subramaniaswamy, V.; Ganesan, M.; Namachivayam, R.K. IIoT for Smart Cities. In *Deep Learning and Blockchain Technology for Smart and Sustainable Cities*; Auerbach Publications: New York, NY, USA, 2025; p. 154.
6. Altulaihan, E.; Almaiah, M.A.; Aljughaiman, A. Cybersecurity threats, countermeasures and mitigation techniques on the IoT: Future research directions. *Electronics* **2022**, *11*, 3330. [\[CrossRef\]](#)
7. Lone, A.N.; Mustajab, S.; Alam, M. A comprehensive study on cybersecurity challenges and opportunities in the IoT world. *Secur. Priv.* **2023**, *6*, e318. [\[CrossRef\]](#)
8. Dong, S.; Shu, L.; Xia, Q.; Kamruzzaman, J.; Xia, Y.; Peng, T. Device identification method for internet of things based on spatial-temporal feature residuals. *IEEE Trans. Serv. Comput.* **2024**, *17*, 3400–3416. [\[CrossRef\]](#)
9. Thakur, H.N.; Al Hayajneh, A.; Thakur, K.; Kamruzzaman, A.; Ali, M.L. A Comprehensive Review of Wireless Security Protocols and Encryption Applications. In Proceedings of the 2023 IEEE World AI IoT Congress (AIoT), Seattle, WA, USA, 7–10 June 2023; IEEE: Seattle, WA, USA, 2023; pp. 0373–0379.
10. Kwon, H.Y.; Kim, T.; Lee, M.K. Advanced intrusion detection combining signature-based and behavior-based detection methods. *Electronics* **2022**, *11*, 867. [\[CrossRef\]](#)
11. Butt, U.J. Developing a Usable Security Approach for User Awareness Against Ransomware. Ph.D. Thesis, Brunel University London, Uxbridge, UK, 2023.
12. Tanikonda, A.; Pandey, B.K.; Peddinti, S.R.; Katragadda, S.R. Advanced AI-Driven Cybersecurity Solutions for Proactive Threat Detection and Response in Complex Ecosystems. *J. Sci. Technol.* **2022**, *3*, 196–218. [\[CrossRef\]](#)
13. Roy, A.; Dhar, A.; Tinny, S.S. Strengthening IoT Cybersecurity with Zero Trust Architecture: A Comprehensive Review. *J. Comput. Sci. Inf. Technol.* **2024**, *1*, 25–50.
14. Zanasi, C.; Russo, S.; Colajanni, M. Flexible zero trust architecture for the cybersecurity of industrial IoT infrastructures. *Ad Hoc Netw.* **2024**, *156*, 103414. [\[CrossRef\]](#)
15. Wang, Y.; Wang, Y.; Liu, J.; Huang, Z.; Xie, P. A survey of game theoretic methods for cyber security. In Proceedings of the 2016 IEEE First International Conference on Data Science in Cyberspace (DSC), Changsha, China, 13–16 June 2016; IEEE: Changsha, China, 2016; pp. 631–636.
16. Ogunbodede, O.O. Game Theory Classification in Cybersecurity: A Survey. *Appl. Comput. Eng.* **2023**, *2*, 669–677. [\[CrossRef\]](#)
17. Messabih, H.; Kerrache, C.A.; Cheriguene, Y.; Calafate, C.T.; Bousbaa, F.Z. An Overview of Game Theory Approaches for Mobile Ad-Hoc Network’s Security. *IEEE Access* **2023**, *11*, 107581–107604. [\[CrossRef\]](#)
18. Butt, M.A.; Ajmal, Z.; Khan, Z.I.; Idrees, M.; Javed, Y. An in-depth survey of bypassing buffer overflow mitigation techniques. *Appl. Sci.* **2022**, *12*, 6702. [\[CrossRef\]](#)

19. Shahriar, M.H.; Khalil, A.A.; Rahman, M.A.; Manshaei, M.H.; Chen, D. iattackgen: Generative synthesis of false data injection attacks in cyber-physical systems. In Proceedings of the 2021 IEEE Conference on Communications and Network Security (CNS), Tempe, AZ, USA, 4–6 October 2021; IEEE: Tempe, AZ, USA, 2021; pp. 200–208.
20. Fereidouni, H.; Fadeitcheva, O.; Zalai, M. IoT and man-in-the-middle attacks. *Secur. Priv.* **2025**, *8*, e70016. [[CrossRef](#)]
21. Ali, M.H.; Jaber, M.M.; Abd, S.K.; Rehman, A.; Awan, M.J.; Damaševičius, R.; Bahaj, S.A. Threat analysis and distributed denial of service (DDoS) attack recognition in the internet of things (IoT). *Electronics* **2022**, *11*, 494. [[CrossRef](#)]
22. Shukla, P.; An, L.; Chakraborty, A.; Duel-Hallen, A. A robust Stackelberg game for cyber-security investment in networked control systems. *IEEE Trans. Control Syst. Technol.* **2022**, *31*, 856–871. [[CrossRef](#)]
23. Zhang, Y.; Malacaria, P. Bayesian Stackelberg games for cyber-security decision support. *Decis. Support Syst.* **2021**, *148*, 113599. [[CrossRef](#)]
24. Ait Temghart, A.; Marwan, M.; Baslam, M. Stackelberg security game for optimizing cybersecurity decisions in cloud computing. *Secur. Commun. Netw.* **2023**, *2023*, 2811038. [[CrossRef](#)]
25. Wang, Z.; Shen, H.; Zhang, H.; Gao, S.; Yan, H. Optimal DoS attack strategy for cyber-physical systems: A Stackelberg game-theoretical approach. *Inf. Sci.* **2023**, *642*, 119134. [[CrossRef](#)]
26. Clempner, J.B. Learning Deceptive Tactics for Defense and Attack in Bayesian–Markov Stackelberg Security Games. *Math. Comput. Appl.* **2025**, *30*, 29. [[CrossRef](#)]
27. Xu, S.; Guan, Y.; Shen, Y. A Stackelberg game for optimal control energy of multi-agent networks under cyber-attacks. *Int. J. Control* **2025**, 1–12. [[CrossRef](#)]
28. Bertsekas, D.P. Nonlinear programming. *J. Oper. Res. Soc.* **1997**, *48*, 334–334. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.