SpacSec 2026 | Technical Track

Adaptive Distributed Denial-of-Service Defense Using Beh. Analysis and Machine Learning

Authors:	M.V. Mitrofanov, N.N. Kramskoy, A.S. Kurakin
Affiliations:	ITMO University & Special Technology Center

Executive Summary

+21.4% F1-Score | 10x FPR Reduction

Integrating Mahalanobis distance behavioral profiling with ensemble machine learning enables adaptiv, adaptive, multi-level DDoS mitigation, achieving sub-200ms detection latency while maintaining 99.98% service availability against dynamic botnet vectors.

The Stricter is Better Fallacy

Target Sectors (Q3 2025):
- Fintech: **26.1%**
- E-commerce: **22.0%**
- Media: **15.8%**

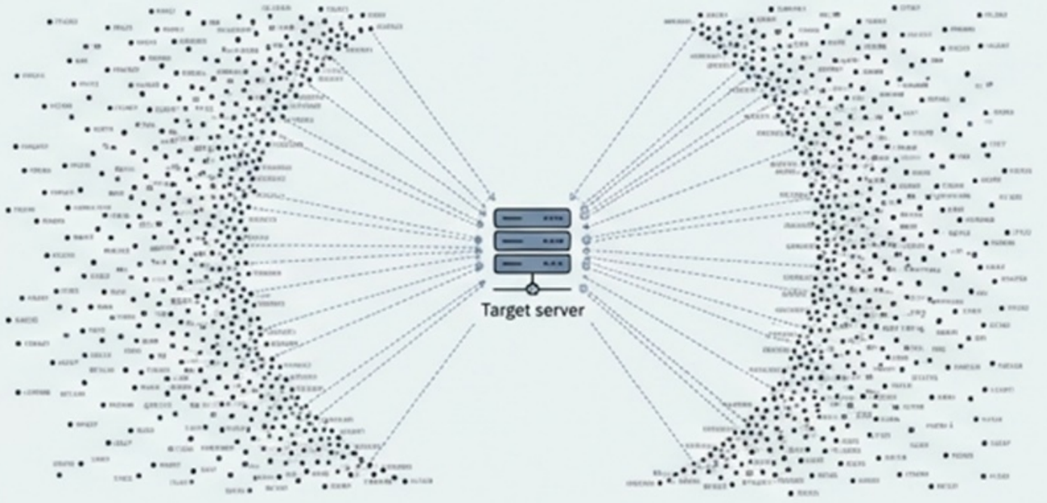
Mechanism: **Aggressive static rate-limiting** causes mass **deduplication failures**, dropping legitimate requests during peak hours and degrading service quality.

Evolution to Low-and-Slow Attack Vectors

`Parameter:` Rate
`Value:` `1-2 packets/s per host`

`Parameter:` Duration
`Value:` `5 minutes per hour per IP`

`Impact:` Operates below **traditional static thresholds**; easily bypasses perimeter caches and signature-based WAF/NGFW layers



Failure of the Static Baseline

Traditional signature-based 10 Gbit/s fixed-threshold systems rely on rigid logic. They lack real-time adaptability to differentiate between legitimate surges and synchronized low-intensity volumetric floods.

System Variable Formalization

Observation Set: `X = {x_1, …, x_n} over time t in [0,T]`

Target Labels: `y_i in {0, 1, 2, 3}`

Dual-Optimization Mathematical Objective

Formula [Eq. 1 - F1-Score]:

$$\mathrm{F1} = \frac{2 * P * R}{P + R}$$

Formula [Eq. 9 - Weighted Loss]:

$$L(\theta) = \alpha * \mathrm{CCE}(y_hat, y) + \beta * \mathrm{FPR}(y_hat, y) + (1 - \alpha - \beta) * \mathrm{FNR}(y_hat, y)$$

Goal: Maximize F1 (Eq. 1) while balancing the loss function (Eq. 9) to heavily penalize both false alarms (FPR) and missed attacks (FNR).

y=0 | Legitimate

Profile: Normal user interactions, expected load spikes.

Target OSI: N/A

y=1 | Volumetric

Profile: Bandwidth consumption avalanches.

Vectors: UDP streams, ICMP storms, DNS amplification.

Target OSI: L3 / L4

y=2 | Protocol

Profile: Exploitation of network interaction state limits.

Vectors: SYN floods (exhausting half-open connection tables).

Target OSI: L4

y=3 | Application

Profile: Resource-intensive logic abuse with minimal channel load.

Vectors: GraphQL inspection (1-10k req/s), WebSocket floods (100k conn/s), gRPC ping-pong, Login spray.

Target OSI: L7

Step 1: Feature Mapping & Z-Score Normalization

Feature Vector [Eq 2]: $\Phi(x_i) = [f_{vol}, f_{prot}, f_{app}, f_{temp}]^T$

Standardization [Eq 10]: $\tilde{\Phi}_j(x_i) = \left(\Phi_j(x_i) - \mu_{train,j}\right) / \sigma_{train,j}$

Step 2: Ledoit-Wolf Shrinkage Regularization

Purpose: Prevents singular matrices in highly correlated feature dimensions.

Formula [Eq 4]: $\Sigma_t = (1-\gamma)\Sigma_t + \gamma * \frac{tr(\Sigma_t)}{d} * I$

Step 3: Covariance-Aware Distance Computation

Formula [Eq 5]: $A(x_i) = [\Phi(x_i) - \mu_t]^T * \Sigma_t^{-1} * [\Phi(x_i) - \mu_t]$

Insight: The squared Mahalanobis distance detects combinatorial anomalies even when individual metrics remain within expected ranges.

Classifier Architecture

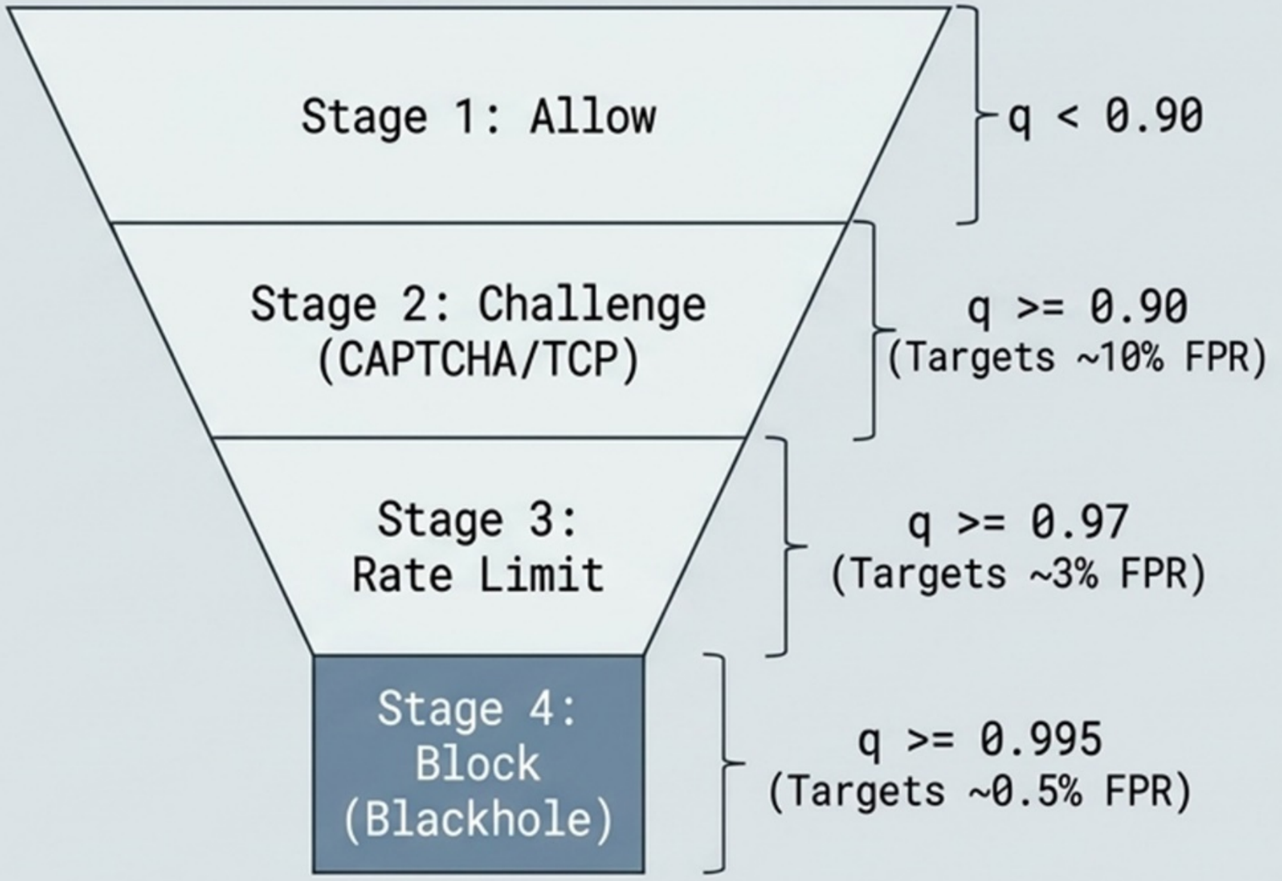
Model: Random Forest Ensemble

- Trees: 500
- Max Depth: 15
- Min Samples Leaf: 5

Loss Function [Eq 8]:
Multiclass Categorical Cross-Entropy (CCE)

Action: Classifies target y_i into the 4 taxonomy vectors.

Adaptive Quantile Mitigation (Eq 6 & 7)



Testbed Parameters

Compute: Intel Xeon Gold 6338 (32 cores)

Memory: 256 GB RAM

Acceleration: NVIDIA A100 GPU
Frameworks: PyTorch 2.1, scikit-learn 1.3.2, Scapy 2.5

Execution: 10 independent runs with distinct random seeds.

Stratified Evaluation Corpus

Datasets:
- CIC-DDoS2019: 5.23M flows
- Bot-IoT: 2.9M flows
- Synthetic low-and-slow: 48k flows

Class Distribution:
- 61.8% Legitimate
- 16.4% Volumetric
- 13.9% Protocol
- 7.9% Application

Validation Protocol: 70% Train / 15% Validation / 15% Test.

Leakage Prevention: Split strictly by capture day; deduplication of retransmitted signatures.

Classification Efficacy vs. Baseline

F1-Score: 0.940 ± 0.008
(+21.4% gain over 10 Gbit/s static baseline)

FPR: 0.52 ± 0.06% (vs 5.12% baseline)

Operational Overhead

End-to-End Latency: 185 ± 9 ms
(Compliant with sub-200ms requirement)

Legitimate Traffic Penalty:
3.2 ms incremental overhead
(Compliant with <5ms threshold)

Detection Accuracy by Taxonomy

Simplifion matrix highlights:

True Negative: 97.8% accurate identification of Legitimate Traffic.

True Positive (Volumetric): 96.1% recall.

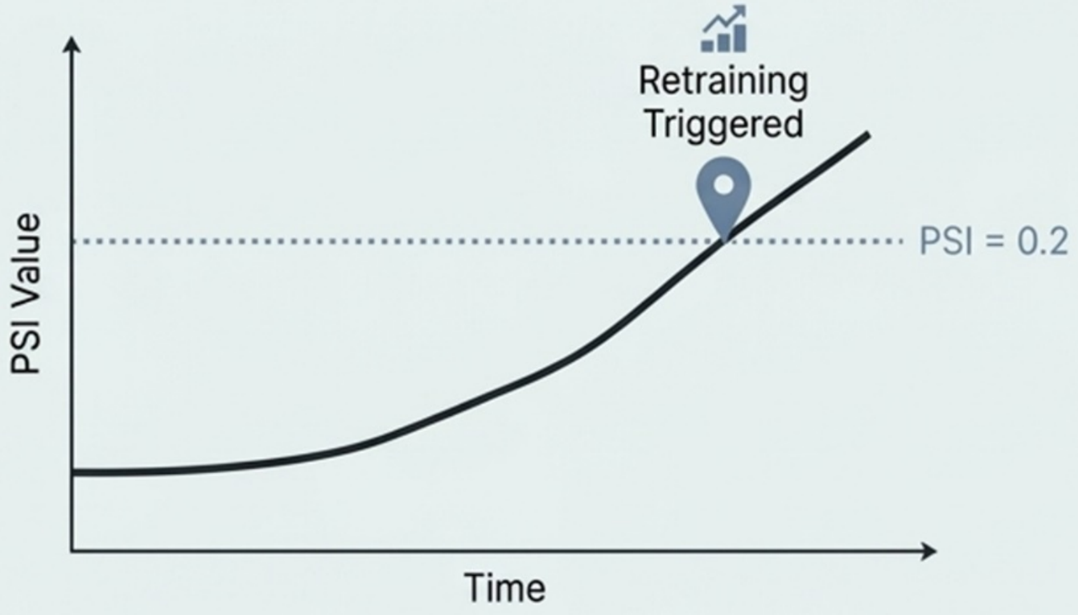
True Positive (Protocol): 95.6% recall.

True Positive (Application): 94.5% recall.

Population Stability Index (PSI) Monitoring

Threshold: Retraining triggered if PSI > 0.2 across 10 histogram bins.

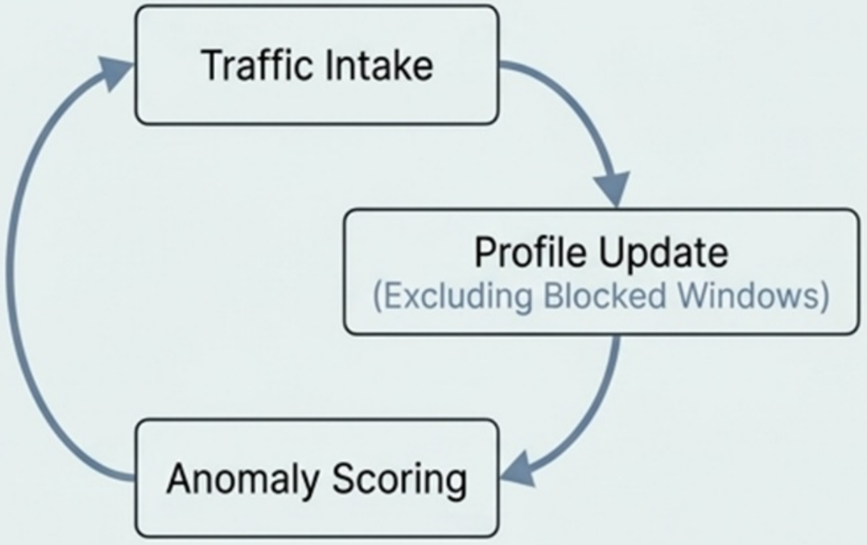
Metric: Median inter-retraining time observed in 30-day trial: 5.4 hours.



Safe Incremental Retraining

Safety Bound: Forced retrain at >= 6-hour elapsed time.

Contamination Resistance: The system strictly excludes time windows flagged as Rate-Limit or Block from the next legitimate profile update. This prevents slow poisoning of the behavioral baseline by undetected persistent attacks.



Benchmark vs. Recent Literature

	SDN + LSTM	DDoS-MSCT	FTG-Net-E	Proposed Solution
Focus	Autoencoder	CNN + Transformer	Hierarchical GNN	Mahalanobis + ML Ensemble
F1-Score	99.1%	99.95%	99.29%	94.0% (Multi-dataset)
Response Type	Binary / None	Binary / None	Binary / None	4-Level Dynamic Policy
Adaptation	Manual Retrain	Offline Update	Offline Update	Automated PSI-Trigger

The Methodological Gap

Insight: While deep-learning detectors show marginally higher isolated benchmark F1 scores, they output binary decisions without integrated response logic. The proposed solution is the only methodology actively linking anomaly scores directly to an interpretable, operational **4-level mitigation policy**.

SLA Constraints

Constraint: Availability Shortfall

Target: 99.99% **availability.**

Actual: Measured at 99.98%.

Impact: Represents a gap of 0.01% **(approx.** 4.32 additional minutes of downtime per month) against ideal operational targets.

Adversarial Risks

Constraint: Feature-Space Vulnerabilities

Limitation: Machine **learning classifiers** remain **theoretically vulnerable** to Fast Gradient Sign Method (FGSM) or **crafted perturbations.** Adversarial-robustness benchmarking is required for next-phase deployment.

Hardware Limits

Constraint: Line-Rate Scaling

Limitation: Benchmarks achieved on A100 GPU **clusters.** Scaling to multi-100-Gbit/s **edge line-rates** will strictly necessitate **offloading inference** to programmable data planes or FPGA accelerators.

Methodological Conclusions

1. F1-score elevated by **21.4%** via combined taxonomy and behavioral profiling.
2. Execution maintained within **sub-200ms** latency bounds.
3. Concept drift successfully managed via **PSI-triggered**, contamination-resistant updates.

Development Roadmap

Vector 1: Federated Learning

Goal: Privacy-preserving distributed defense models.

Vector 2: Reinforcement Learning

Goal: Dynamic attacker-adaptation policies.

Vector 3: Explainable AI (XAI)

Goal: Increasing SOC operator trust and regulatory compliance.

Open for Technical Review

SpacSec 2026 Floor Discussion & Data Access Inquiries