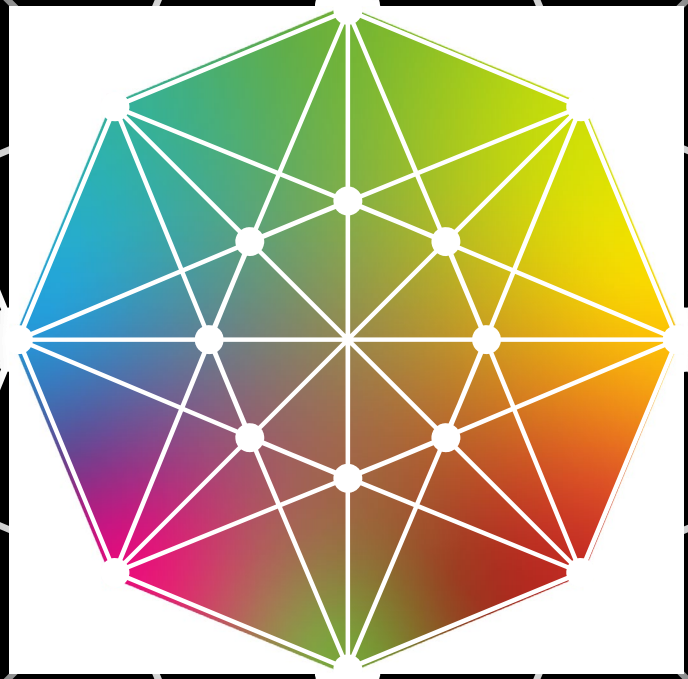
Resilience Against Adversarial Concept Drift in Port Scanning

Batch vs. Incremental Paradigms for Green AI

7-9 July 2026 • Geneva, Switzerland



ITU KALEIDOSCOPE
GENEVA2026

Piero Aliaga & **Jose Menendez**

Universidad Nacional de Ingenieria (UNI), Lima, Peru

Universidad Nacional de Ingenieria (UNI), Lima, Peru

1 The Challenge: Adversarial Concept Drift & The "Window of Exposure"

ITU KALEIDOSCOPE GENEVA2026

Port scanning is the reconnaissance phase of most APTs.

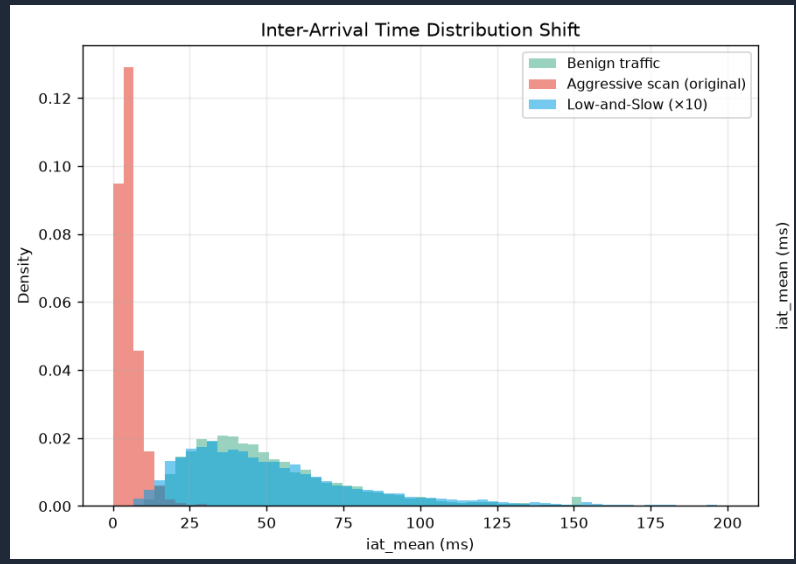
Static, anomaly-based NIDS learn a fixed decision boundary from historical traffic.

"Low-and-Slow" attackers deliberately reshape their traffic to statistically mimic benign behaviour, evading detection thresholds entirely.

Research Question

How does a Model-Agnostic Adaptive Defense Architecture relate to IDS resilience and operational efficiency against adversarial drift?

Same scan, different statistical signature

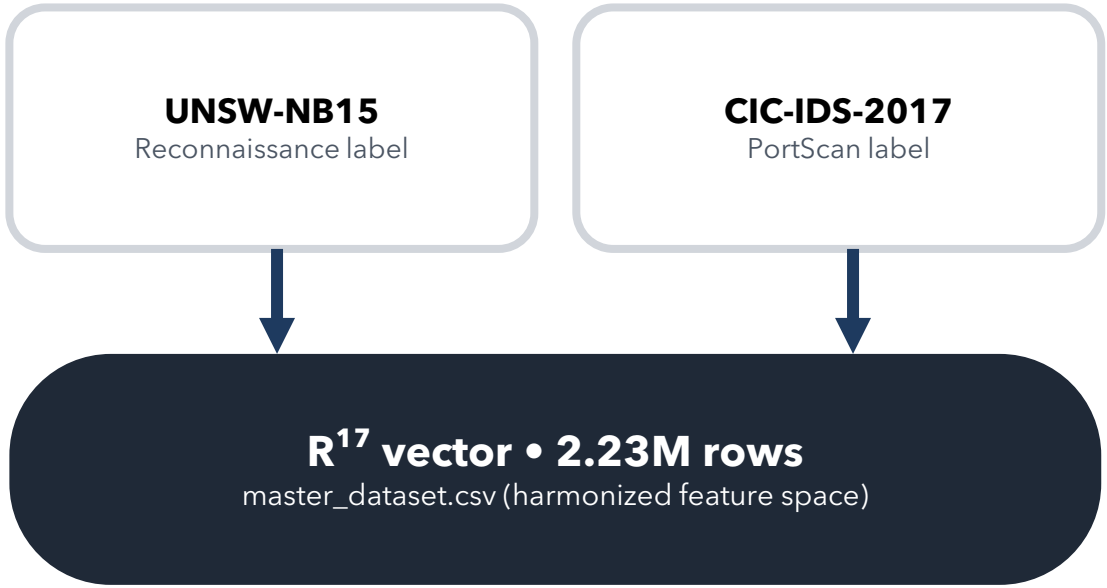


"Adversarial AI violates the i.i.d. assumption: $P_{test}(x,y) \neq P_{train}(x,y)$"

2 Methodology: Data & Adversarial Attack

ITU KALEIDOSCOPE GENEVA2026

Phase 1 – Data Harmonization



Key feature: **iat_mean** (mean inter-arrival time) The single most discriminative signal, and the one the attacker deliberately targets.

The "Super Evasive" Attack

Eq. 8 (Statistical Mimicry)

$$x'_{vol} = \mu_{benign} + \varepsilon$$

Volumetric features sampled from the benign distribution

Eq. 9 (Low-Rate Evasion)

$$x'_{time} = x_{time} \times 10$$

iat_mean dilated 10× (the "Low-and-Slow" signature)

Result: malicious flows that statistically resemble benign traffic to a static classifier, the "Super Evasive" dataset.

3 A Model-Agnostic Adaptive Agent

ITU KALEIDOSCOPE GENEVA2026

The "Brain" (Watchdog) and the "Muscle" (Classifier) evolve independently

Traffic Stream

Batches B_1, B_2, ... B_n

WATCHDOG

$D_n = \sup_x |F_{ref}(x) - F_{act}(x)|$

$p < 0.01$

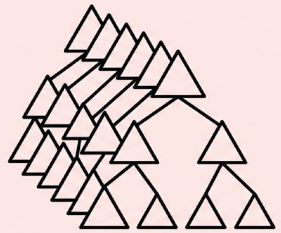
Drift!

Paradigm A: Random Forest

Full retraining • O(N)

2.23M rows → 356.57s

$O_{RF} = O(M \cdot N \log N)$

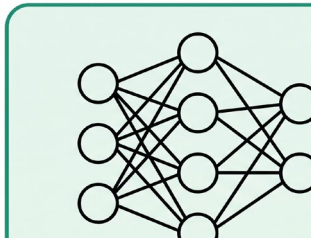


Paradigm B: CNN-LSTM

Fine-tuning • O(B)

2,000 rows → 2.61s

$O_{CNN\text{-}LSTM} = O(B \cdot P)$



The Watchdog operates **independently of the classifier** drift detection is decoupled from classification logic, enabling modular, swappable security architectures.

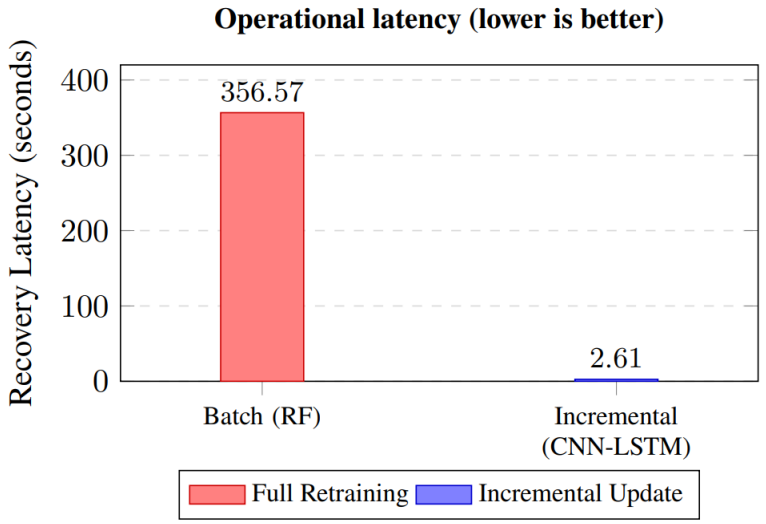
4 Results: A Silent Collapse, A Costly Recovery

ITU KALEIDOSCOPE GENEVA2026

Metric	RF (Batch)	CNN-LSTM (Incr.)
Scenario 1 — Pure Evasion (adversarial only)		
Precision	1.0000	1.0000
Recall	0.0291	0.0428
F1-Score	0.0566	0.0820
Scenario 2 — Realistic Mix (benign + adversarial)		
Precision	1.0000	0.9852
Recall	0.0291	0.0428
FPR	0.0000	0.0006
ROC-AUC	**0.9906**	**0.6444**

Table 3 – mixed evaluation set. Neither model raises a false alarm (FPR ≤ 0.06%).

Time-to-Recovery (lower is better)



Threshold Misalignment Under Distribution Shift

0.99 AUC + **2.9%** Recall

RF's ranking stays **nearly perfect**, but the post-drift distribution no longer aligns with the static 0.5 threshold. "Threshold misalignment," not a ranking failure.

Both paradigms recover, **only one** in operational time.

136.5×

speedup in Time-to-Recovery

The Window of Exposure

On a 100 Gbps link, a 6-minute blind spot allows for **~4.5 TB** of uninspected exfiltration.

5 Conclusion: Plasticity as Security

ITU KALEIDOSCOPE GENEVA2026

Differentiability, the ability to update weights via gradients is **not optional. It is a non-negotiable requirement** for real-time autonomous defense.

Green AI

O(B) fine-tuning vs. O(N) retraining efficiency that scales independently of historical data size, enabling deployment on Edge IoT gateways.

Acknowledged Limits

Catastrophic forgetting on D0 not empirically measured. Univariate K-S watchdog only monitors iat_mean.

Future Work

Elastic Weight Consolidation (EWC), multivariate MMD drift detection, and validation via real red-team exercises.

Thank you!

Piero Aliaga & Jose Menendez
Universidad Nacional de Ingenieria, Lima, Peru