Italian National Research Council

ICAR CNR

From One to Many: Few-Shot Deep Ensembles for Slow DoS Attack Detection

Alberto Falcone, Massimo Guarascio, Angelica Liguori, Francesco Sergio Pisani, Francesco Scala

Institute for High Performance Computing and Networking

Italian National Research Council (ICAR – CNR)

ECML PKDD 2025 PORTO, PORTUGAL

MLCS 2025

Problem Definition

Problem

- Slow DoS attacks mimic legitimate/normal traffic, making them difficult to detect
- Few labeled samples available for supervised methods
- Unsupervised approaches often prone to false positives

Goal

- Detect slow DoS attacks efficiently
- Operate under **few-shot** and **imbalanced data** settings

Problem Definition

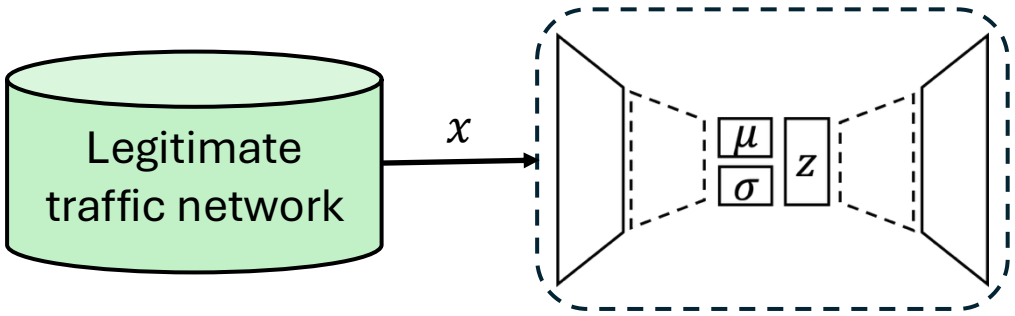
Idea

- Use **Autoencoder-based architectures** to model normal traffic
- Instantiate multiple detectors with different sensitivity thresholds
- Combine their outputs using an ensemble strategy

Solution

- MoAE^2 (Mixture of AutoEncoder Experts), a semi-supervised framework that leverages one autoencoder to build multiple threshold-based detectors and combines them through a Mixture-of-Experts trained with only a few labeled samples

Methodology Overview

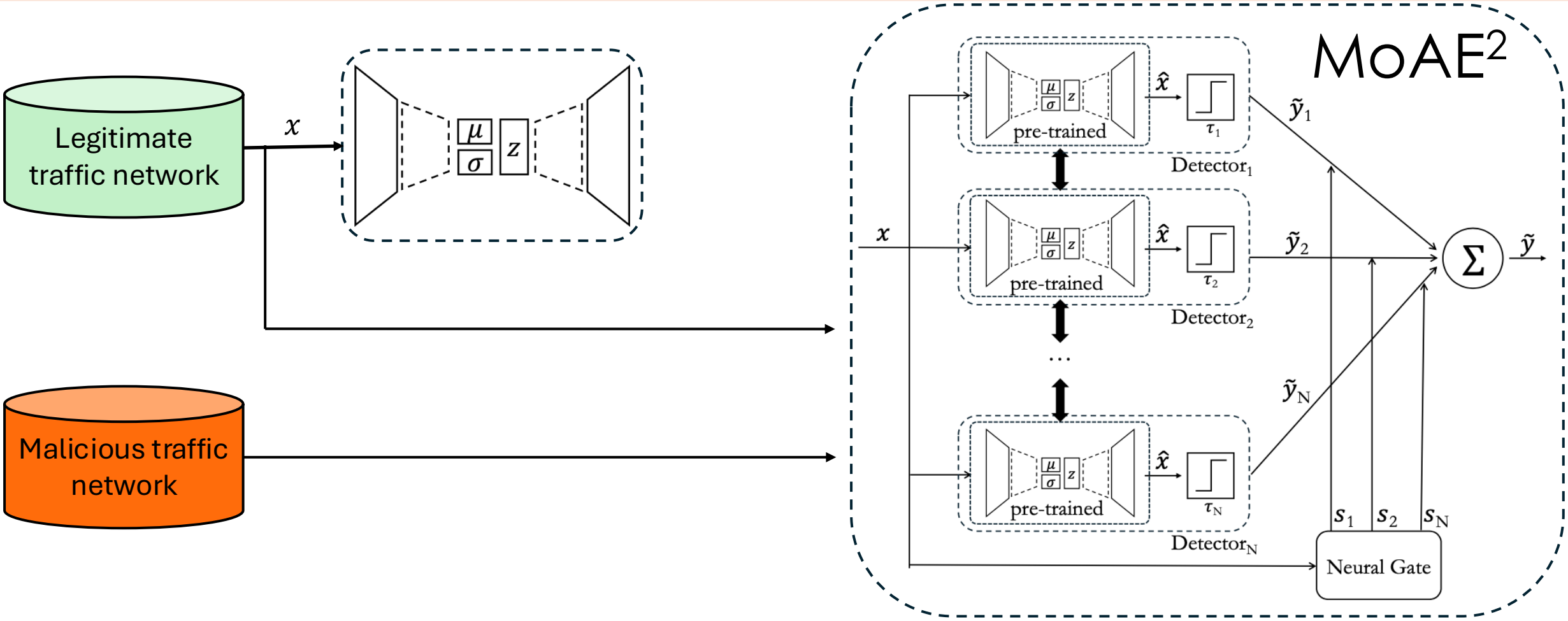


Methodology Overview

Phase 1: Unsupervised learning

- Train a **Variational Autoencoder** (VAE) on benign traffic
- Learns the distribution of **normal behavior**
- **Reconstruction error** used as anomaly indicator
  - Anomalous samples → higher reconstruction error
- Apply **threshold-based detection**
  - Any sample with a reconstruction error exceeding a predefined threshold is flagged as anomalous

Methodology Overview



Methodology Overview

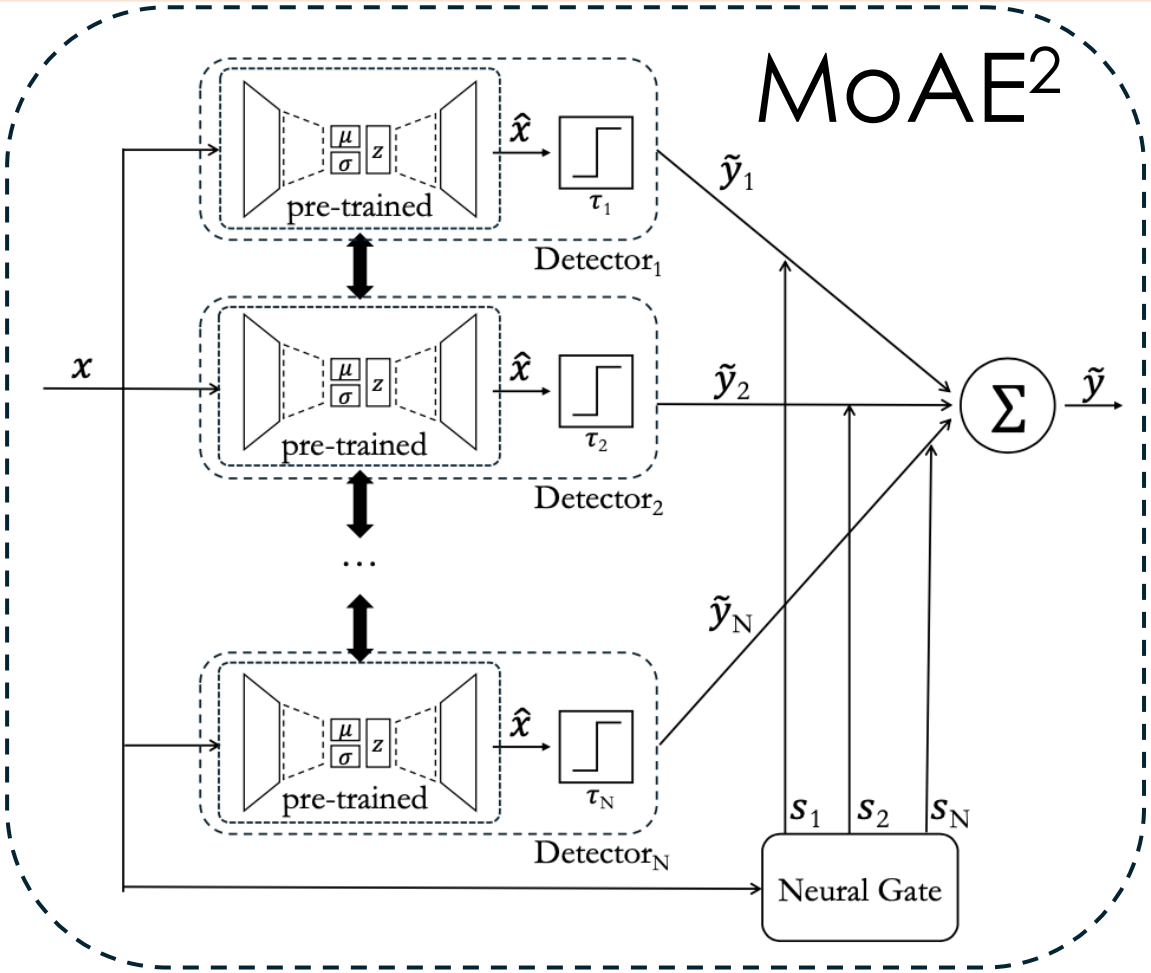
Phase 2: Ensemble & Few-shot learning

- Apply **different thresholds** to reconstruction error, producing different **base detectors**
  - Thresholds are chosen from the error distribution
- Each detector has a different sensitivity to anomalies
  - Each produces a different prediction
- The predictions are combined with a Mixture-of-Experts

Methodology Overview

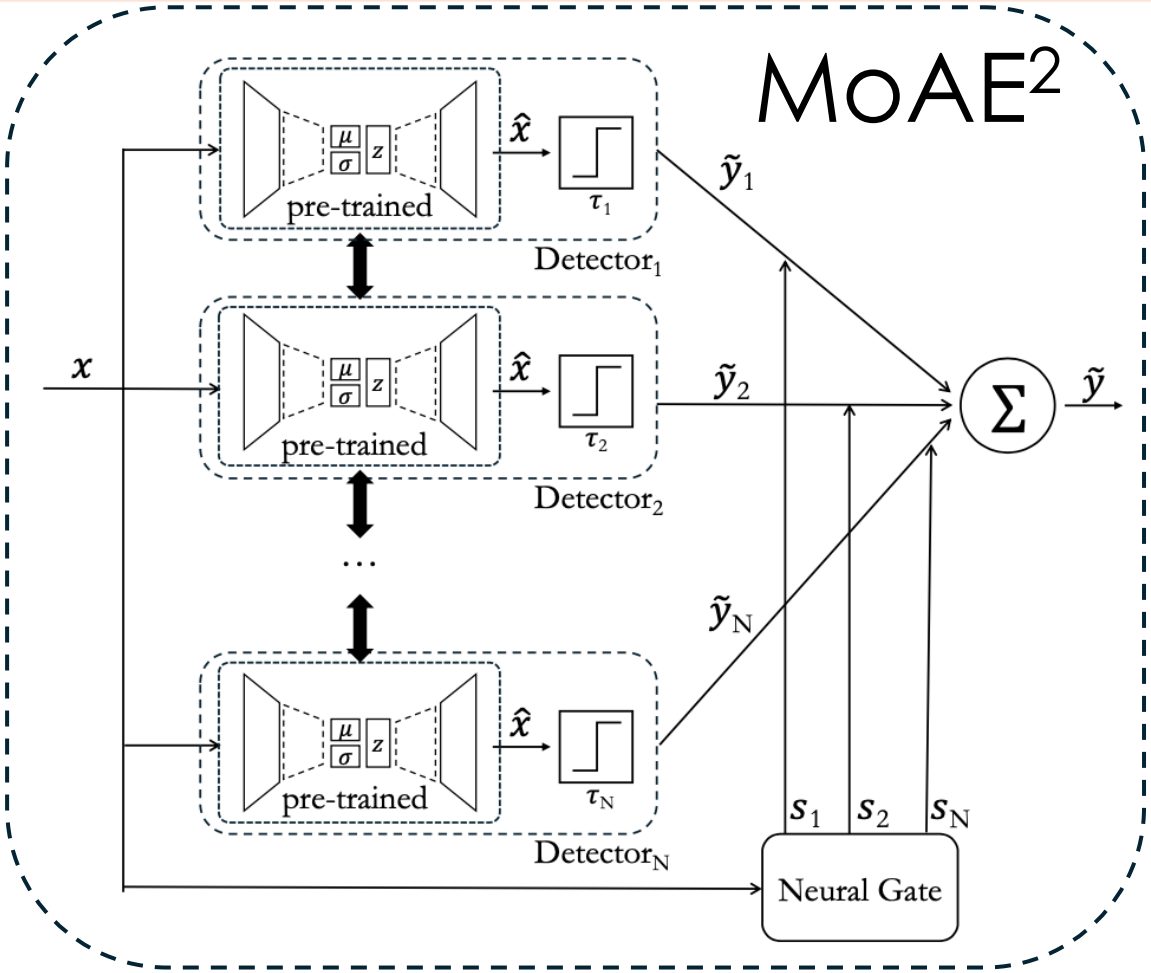
- A **neural gating network** assigns weights to the detectors
- Final prediction is a weighted combination of detector outputs

$$\tilde{y} = \sum_{i=1}^{N} \tilde{y}_i \cdot s_i$$



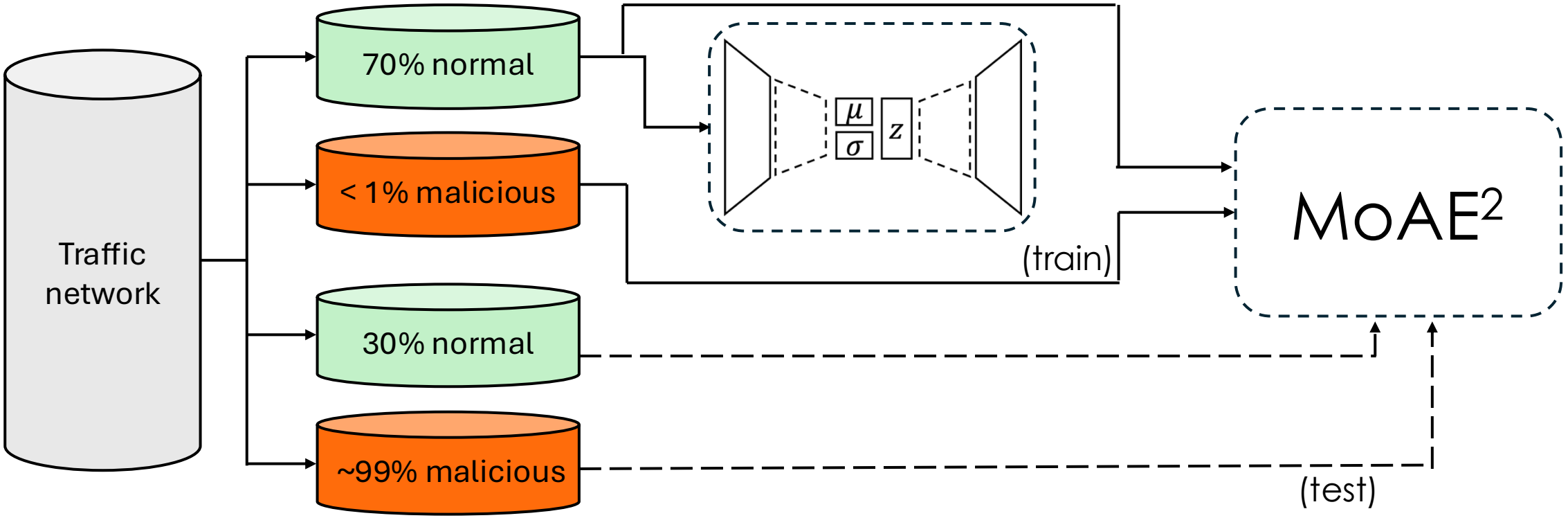
Methodology Overview

- The neural gate is trained within a **few-shot learning framework**
  - By using both legitimate traffic and a limited number of anomalies
- Base detectors remain **frozen**, providing only their predictions



Experimental Evaluation

- MoAE2 is tested on a well-known benchmark dataset: the **CIC-IDS2017**, containing both legitimate and malicious traffic



Experimental Evaluation

- MoAE² outperforms baselines
- +25% F1-score and +21% GMean over XGBoost (best baseline)

Technique	F1-Score	GMean	Δ-F1%	Δ-GMean%
Decision Tree	0.22	0.45	- %	- %
KNN	0.28	0.48	27.3%	6.7%
Random Forest	0.27	0.48	22.7%	6.7%
XGBoost	0.36	0.53	63.6%	17.8%
MoAE²	**0.45**	**0.64**	**104.5%**	**42.2%**

Conclusion and Future directions

- MoAE², a semi-supervised approach for detecting slow DoS attacks, is proposed
- Experiments demonstrate the effectiveness of MoAE² compared to ML-based baselines

- As future work, we plan to increase diversity among base detectors and improve the efficiency of the MoE strategy

Thank You for your attention!

Questions?

Angelica Liguori
ICAR-CNR
angelica.liguori@icar.cnr.it