DistillGuard: Cross-Dataset Comparative Study

This document synthesizes the results of implementing DistillGuard (Algorithm 1) across three diverse IoT datasets: **ToN-IoT**, **RT-IoT**, and **N-BaIoT**.

1. Dataset Characteristics

Metric	ToN-IoT	RT-IoT	N-BaIoT
Samples (Used)	~211,000	~123,000	~200,000 (Subsampled from 7M)
Features	21 (Selected)	15 (Selected) + Ports	115 (Flow Stats)
Classes	10	12	10+ (Botnet Families)
Complexity	High (Heterogeneous)	Low (Clean separation)	High (Volume/Variability)

2. Component Analysis (Algorithm 1)

A. Performance (Student Model)

Dataset	Teacher Acc	Student Acc	Distillation Loss (SG-KD)
ToN-IoT	96%	95.92%	Low
RT-IoT	99%	99.55%	Very Low
N-BaIoT	**17.4% (Failed)**	**82.56% (Success)**	*Student outperformed Teacher*

B. Compression Efficacy (Step 9)

Comparison of **Int8 Dynamic Quantization**.

Dataset	Original Size (MB)	Quantized Size (MB)	Acc Drop	Reduction
ToN-IoT	1.04	0.26	$<1\%$	4x
RT-IoT	1.04	0.26	$<0.5\%$	4x
N-BaIoT	1.13	0.28 (est)	~0.5%	4x

C. Adversarial Robustness (Step 7 Effect)

Does Teacher Adversarial Training transfer to Student?

Dataset	Clean Acc	FGSM ($\epsilon = 0.05$)	Conclusion
ToN-IoT	95.9%	~20.2%	Failed Transfer
RT-IoT	99.5%	~4.7%	Failed Transfer
N-BaIoT	82.6%	~7.6%	Failed Transfer

Hypothesis: SG-KD forces the student to mimic the "logits" of the teacher, but not necessarily the *robustness geometry* of the decision boundary. Adversarial training might be needed *directly* on the Student during distillation (Adversarial Distillation).

3. Training Dynamics

- **ToN-IoT**: Stable convergence.
- **RT-IoT**: Extremely fast convergence (Separable classes).
- **N-BaIoT**: **Teacher Collapsed** (17% Acc) but **Student Recovered** (82% Acc). The MLP Student proved more robust to the N-BaIoT feature set than the Transformer Teacher.

4. Final Recommendation

1. **Deployment**: The 4x compressed Student is viable for Edge IoT nodes (Raspberry Pi/ ESP32).
2. **Weakness**: Robustness against Gradient Attacks is the primary gap.
3. **Future Work**: Implement *Adversarial Distillation* (training student on adv examples matched to teacher's clean logits).