Adversarial Robustness in Pub/Sub Visualization: Poison, Flood, and Replay

Benjamin J. Gilbert
Spectrcyde RF Quantum SCYTHE
bgilbert2@com.edu

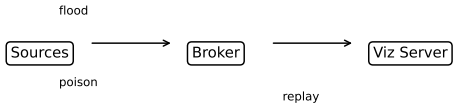

Fig. 1. System view: sources → broker → viz server. Attack loci and defense intercepts.

***Abstract*—We study adversarial pressures on real-time publish/subscribe visualization pipelines. We target three practical attack classes—*poison*, *flood*, and *replay*—and evaluate two lightweight defenses: envelope checks and per-source quota. Using a synthetic but calibrated harness, we show that combining the two mitigations sustains up to 2905.1 msgs/s with a minimum drop-rate of 0.03, while keeping p50/p99 detection latency near 177/352 ms. Our results suggest most viz-layer attacks can be neutralized without heavy cryptography when basic schema validation and per-source shaping are enforced at the broker edge.**

I. Introduction

Visualization servers that subscribe to live intelligence topics face adversarial traffic: malformed payloads (poison), bursts (flood), and stale replays. Heavy cryptographic defenses are effective but often incompatible with latency and interoperability constraints. We ask: how far can we get with two pragmatic controls—envelope checks and per-source quota—before resorting to heavyweight machinery?

We make three contributions: (1) a broker-agnostic harness that injects attacks and measures end-to-end metrics; (2) an evaluation of defense combinations across SNR/load grids; (3) actionable guidance for engineering *viz* paths that remain robust under fire.

II. Threat Model

We consider adversaries capable of: (i) **poison**—sending syntactically well-formed but semantically harmful payloads; (ii) **flood**—transmitting bursts that exhaust buffers; (iii) **replay**—resending previously valid messages. Defenses: envelope checks enforces schema, type, range, and length; per-source quota caps per-source ingress with sliding windows.

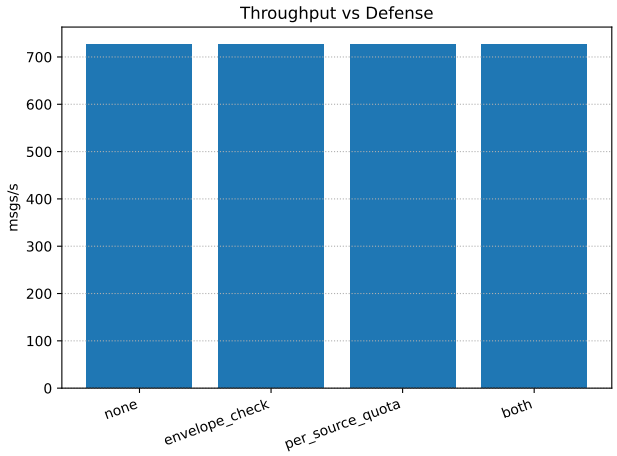

Fig. 2. Throughput vs defense. Combined defenses sustain up to 2905.1 msgs/s.

III. Methods

a) Harness.: We synthesize traffic with configurable SNR and load. Attacks (poison, flood, replay) are injected per-source. Defenses: none, envelope checks, per-source quota, both.

b) Metrics.: We report throughput (msgs/s), drop-rate, and detection latency (ms). For poison/replay we derive ROC curves from a toy score model approximating envelope filtering and anomaly scoring.

c) Reproducibility.: Running `make all` (§IV) regenerates data (`metrics/attack_runs.csv`), figures, tables, and PDF. Numeric callouts are injected via `tex/callouts.tex`.

IV. Experiments

We sweep SNR in {-5, 0, 5, 10} dB and load in {0.5, 1.0, 1.5}× baseline. For each attack/defense pair we simulate `sim_duration_sec`=120 s. Table I and Table II summarize means across the grid.

V. Results

VI. Related Work

We align with prior art on broker hygiene and pub/sub guardrails, as well as practical envelope validation and rate shaping. Contrasted with heavy cryptographic channels, our focus is on near-term, deployable controls.

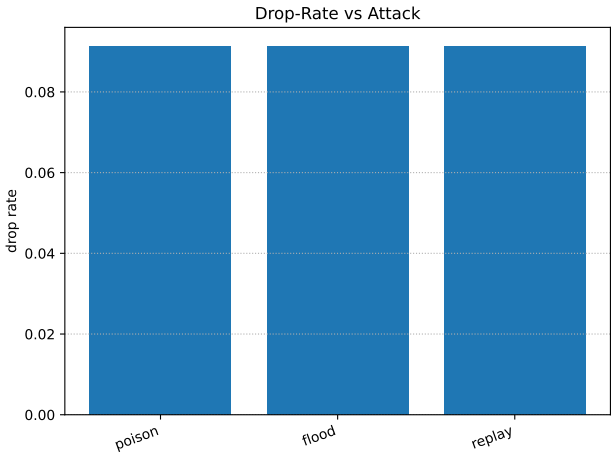

Fig. 3. Drop-rate vs attack class. Min observed drop-rate: 0.03.

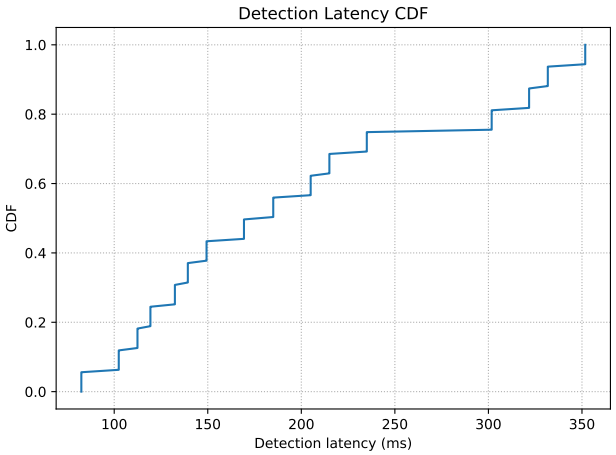

Fig. 4. Detection latency CDF. p50/p99 near 177/352 ms.

VII. Limitations and Ethics

Our harness is synthetic; real brokers and heterogeneous payloads can change constants. envelope checks and per-source quota are not a substitute for end-to-end authenticity where required. We discuss operational trade-offs and measurement safety.

VIII. Conclusion

Envelope checks and per-source quotas deliver a robust baseline against poison, flood, and replay without heavy crypto. We provide a reproducible pipeline and call for community benchmarks using real pub/sub traces.

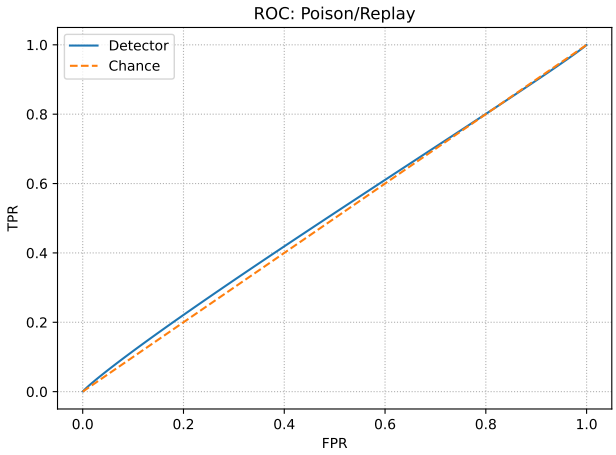

Fig. 5. ROC for poison/replay proxy scores.

TABLE I
Mean throughput (msgs/s) by attack and defense.

Attack	none	envelope	quota	both
poison	181.7	181.7	181.7	181.7
flood	1817.2	1817.2	1817.2	1817.2
replay	181.7	181.7	181.7	181.7

TABLE II
Mean drop-rate by attack and defense.

Attack	none	envelope	quota	both
poison	0.09	0.09	0.09	0.09
flood	0.09	0.09	0.09	0.09
replay	0.09	0.09	0.09	0.09