prmon Memory Anomaly Detection

Warm-up Project

Agnidipto Banik

February 27, 2026

Objective

- Generate memory time-series data using `prmon`
- Inject a controlled memory anomaly using burner tests
- Apply statistical anomaly detection
- Evaluate detection performance quantitatively and visually

Experimental Setup

- Environment: WSL Ubuntu + prmon (built from source)
- Metric used: RSS (Resident Set Size)
- Three segments:
  - Normal baseline
  - Elevated memory usage (anomaly)
  - Recovery to baseline
- Total samples: 228

Anomaly Injection

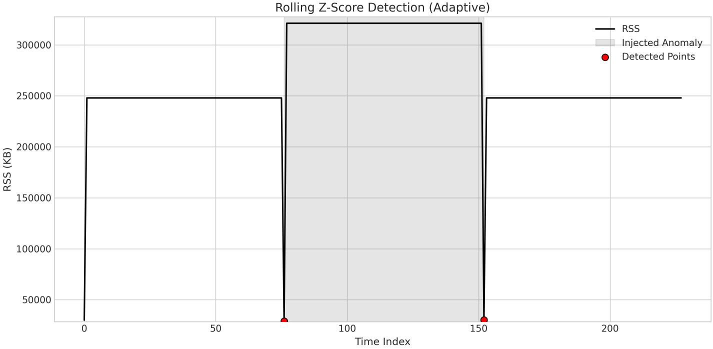
- Normal burner run establishes baseline behavior
- Memory parameters increased to create sustained elevation
- System returned to baseline to simulate recovery
- Anomaly region labeled using known injection interval

Rolling Z-Score (Adaptive Detection)

$$z_t = \frac{x_t - \mu_{window}}{\sigma_{window}}$$

- Compares current value to recent behavior
- Adapts dynamically to new levels
- Threshold: $|z| > 3$

Rolling Z-Score – Results



Why Rolling Fails

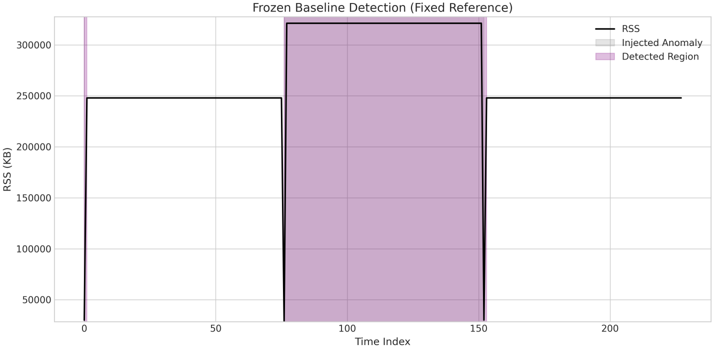
- Detects boundary transitions only
- Rolling window adapts to sustained shift
- Elevated regime becomes statistically normal
- Low recall for step anomalies

Frozen Baseline Z-Score (Fixed Reference)

$$z_t = \frac{x_t - \mu_{baseline}}{\sigma_{baseline}}$$

- Baseline statistics computed from initial normal segment
- No adaptation to new regimes
- Detects sustained structural deviation

Frozen Baseline – Results



Why Frozen Performs Better

- Step anomaly introduces sustained mean shift
- Frozen model remains anchored to healthy baseline
- Entire elevated regime remains statistically significant
- High precision and recall observed

Final Quantitative Results

Rolling Z-Score

- Precision: 0.50
- Recall: 0.013
- F1-score: 0.026
- Confusion Matrix:

$$\begin{bmatrix} 151 & 1 \\ 75 & 1 \end{bmatrix}$$

Frozen Baseline

- Precision: 0.974
- Recall: 1.00
- F1-score: 0.987
- Confusion Matrix:

$$\begin{bmatrix} 150 & 2 \\ 0 & 76 \end{bmatrix}$$

Why Not Machine Learning?

- Warm-up focused on controlled statistical evaluation
- Dataset is small and structured (single step anomaly)
- No need for complex modeling
- Statistical methods are interpretable and sufficient

Why Is the Data Simple?

- The goal was to isolate the behavior of detection methods under a clear regime shift.
- A step increase in memory models realistic scenarios such as memory surges or leaks.
- Keeping the setup controlled makes it easier to interpret why a method succeeds or fails.
- More complex patterns (noise, drift, seasonality) can be explored in later stages.

Trade-offs and Sustainability

- Rolling suitable for short-lived spikes
- Frozen requires clean baseline period
- Frozen may struggle with gradual drift
- Real deployments require adaptive baselines and multi-metric monitoring

Conclusion

- Controlled memory anomaly successfully injected using prmon
- Statistical detection methods applied and evaluated
- Frozen baseline outperformed adaptive rolling for sustained shift
- Demonstrates structured experimental reasoning