IMPERIAL

Anomaly Detection in Streaming Data with Concept Drift Using EVT

Student: Kateryna Kaplunova
Supervisor: Zak Varty

Problem Overview

In modern cloud systems, financial markets, and cyber-defense platforms, vast volumes of data are generated continuously - and the cost of missing or misclassifying anomalies can be catastrophic. Yet, real-time anomaly detection remains a statistically challenging setting: distributions drift, human labeling is rare, and most deployed systems rely on heuristic or score-based methods, offering limited statistical guarantees.

This project addresses a key question in the field:

Can EVT-based methods be extended and reliably applied in streaming anomaly detection with concept drift, and how do they compare to established alternatives?

Background

Extreme Value Theory

EVT describes the probabilistic behavior of extreme events. One of its central tools is the Peaks Over Threshold (POT) approach, which focuses on modeling values that exceed a sufficiently high threshold t.

According to the Pickands–Balkema–de Haan theorem, under broad conditions, the distribution of exceedances converges to the **Generalized Pareto Distribution (GPD)** as t increases:

$$P(X > t + y \mid X > t) \approx GPD(y; \gamma, \sigma)$$

$$GPD(y; \gamma, \sigma) = 1 - \left(1 + \frac{\gamma y}{\sigma}\right)^{-1/\gamma}, \quad y > 0$$

where γ is the shape (tail index), and σ the scale parameter.

Anomaly Detection

Anomaly detection aims to identify rare and unexpected patterns. In the streaming setting, methods must be online (one-pass), memory-efficient and adaptable to non-stationary data.

EVT-Based Streaming Detectors

SPOT (Siffer et al., 2017) uses the POT formulation to set adaptive thresholds for streaming data. SPOT assumes stationarity, makes no assumptions about the full data distribution and continuously updates a GPD fit on the exceedances over a high threshold, allowing it to dynamically estimate a quantile-based decision threshold z_q that maintains a user-defined false positive rate.

DSPOT, also introduced by Siffer et al. (2017), extends this to the non-stationary data by removing local trends via detrending. It computes residuals:

$$X'_t = X_t - M_t$$

where M_t is a local average or trend estimate. The POT method is then applied to X'_t. DSPOT assumes the residuals are i.i.d., satisfying the conditions for EVT.

Preliminary Results

Reported Performance of SPOT. The ROC curve from Siffer et al. (2017), shown below, illustrates how adjusting the risk parameter q allows SPOT to trade off between sensitivity and robustness. This behavior is central to this project, as it highlights how EVT-based methods can be tuned for different operational needs in streaming settings.

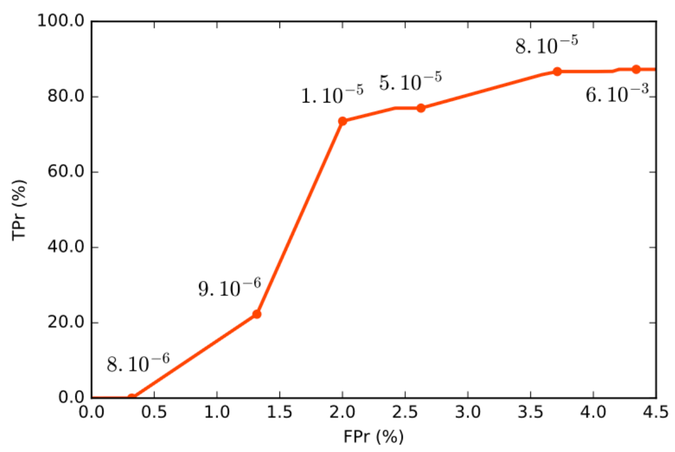

Figure 1: ROC curves for SPOT on MAWI dataset (Siffer et al., 2017).

DSPOT Applied to Real-World Streaming Data. We applied DSPOT to the `occupancy_6005.csv` dataset from the Numenta Anomaly Benchmark - a univariate traffic signal known to exhibit strong periodicity and concept drift. Anomaly labels in NAB are provided as time windows, and a prediction is considered correct if it falls within a labeled window.

As shown below, DSPOT successfully identified the true anomaly window while generating only one false alarm. The model adapts its threshold to the evolving baseline and remains robust to non-stationary behavior.

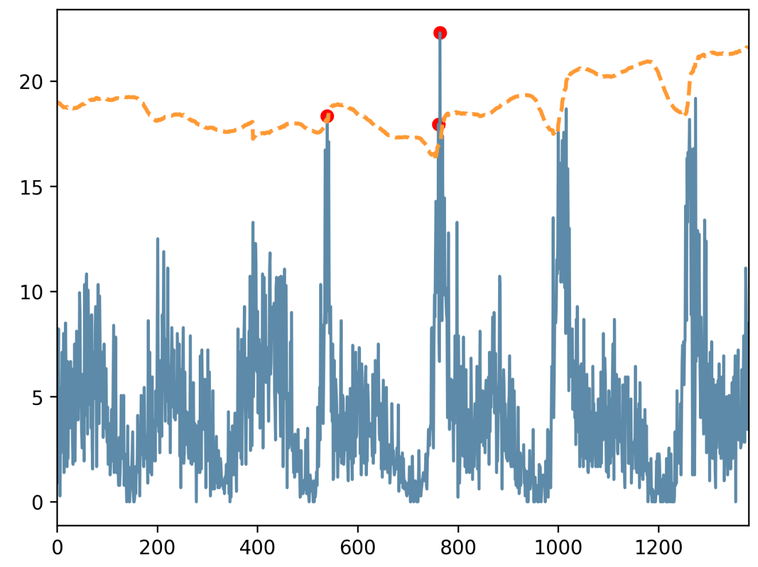

Figure 2: DSPOT detections (red) and adaptive threshold (orange) on `occupancy_6005`.

Project Contributions

The novelty of this work lies in extending DSPOT through:

1. **New context:** First known evaluation of DSPOT on **NAB datasets** - all featuring streaming data with natural concept drift.
2. **Comparative evaluation:** DSPOT is benchmarked against: ADWIN + Isolation Forest, Robust Random Cut Forest and HTM (NAB native).
3. **Extension pathways:**
   (a) **Adaptive window sizing**: Dynamic calibration window selection for better re-adaptation.
   (b) **Combining SPOT with drift detectors**: Use explicit drift detectors (e.g., ADAM) to switch between SPOT and DSPOT modes.
   (c) **EVT on anomaly scores**: Instead of raw data, apply the POT method to the anomaly scores produced by benchmark detectors (e.g., Isolation Forest).

References

1. Siffer, A. et al. (2017). Anomaly Detection in Streams with Extreme Value Theory. KDD.
2. Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values.
3. Numenta Anomaly Benchmark (NAB): https://numenta.org/nab
4. Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey.
5. Sadik, M., and Gruenwald, L. (2014). Research issues in outlier detection for data streams.