NEW YORK INSTITUTE OF TECHNOLOGY

University of Manitoba

Detecting Cyber Anomalies in IoT Traffic Using a Dirichlet–Mixture Bayesian Model

Surani Matharaarachchi [1] Saman Muthukumarana [2]

[1]New York Institue of Technology [2]University of Manitoba

Abstract

This work compares two Bayesian models for detecting right-tail anomalies in **seasonal, over-dispersed** IoT count data. A single-regime Negative Binomial (NB) with a Fourier seasonal basis is contrasted with a Dirichlet-weighted NB mixture modeling *normal* and *spike* regimes. The mixture achieves higher AUC and better-calibrated spike probabilities on synthetic heavy-tailed data.

Problem

Goal: Flag transient, multiplicative spikes in hourly counts (e.g., packets, connections) under seasonality and over-dispersion.

Challenges: Seasonal, daily periodicity, rare heavy-tailed shocks that violate single-regime assumptions.

Data & Simulation

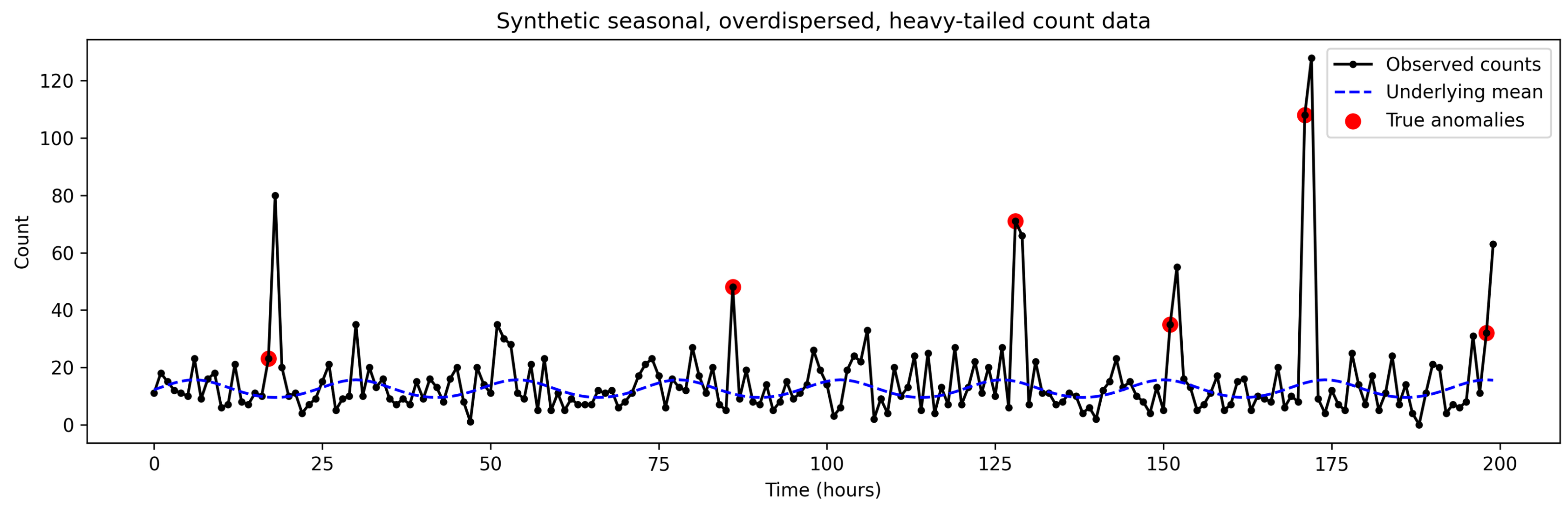
Time horizon $T = 200$ hours with a **daily sinusoid** (period 24) and **no trend**. Underlying mean

$$\mu_t = \exp\Big(2.5 + 0.25\sin\big(2\pi t/24\big)\Big).$$

Counts follow NB2:

$$Y_t \sim \mathrm{NB}\big(\mu_t, \alpha\big), \quad \alpha = 5.$$

We inject 6 random indices with multiplicative, heavy-tailed boosts for 1–2 hours ⇒ ground-truth anomalies.



Model A - Negative Binomial

Seasonality via a small Fourier basis $\mathbf{B}(t)$ (sin/cos of 24h harmonics):

$$\log \mu_t = a_0 + \mathbf{B}(t)^\top \gamma, \quad Y_t \sim \mathrm{NB}(\mu_t, \alpha), \quad \alpha \sim \mathrm{Gamma}(30, 0.3).$$

Inference: ADVI in PyMC.
Score (right-tail): $S_t = 1 - \Pr\big(\tilde{Y}_t \geq y_t \mid \text{posterior}\big)$.
Flag rule: p-right < 0.001.

Model B - Dirichlet Mixture NB

Mixture with shared dispersion and seasonal mean:

$$Y_t \sim \pi_0\,\mathrm{NB}(\mu_t, \alpha) + \pi_1\,\mathrm{NB}(k\mu_t, \alpha)$$

$$\pi \sim \mathrm{Dirichlet}(98, 2), \;\; \log k \sim \mathcal{N}(\log 15,\ 0.4^2).$$

Posterior responsibility for the spike component:

$$r_t = \Pr(\text{spike} \mid y_t, \theta) = \frac{\pi_1\, p(y_t \mid k\mu_t, \alpha)}{\pi_0\, p(y_t \mid \mu_t, \alpha) + \pi_1\, p(y_t \mid k\mu_t, \alpha)}.$$

Score: $S_t = \mathbb{E}[r_t]$. **Flag rule:** $S_t > 0.9$.

Results

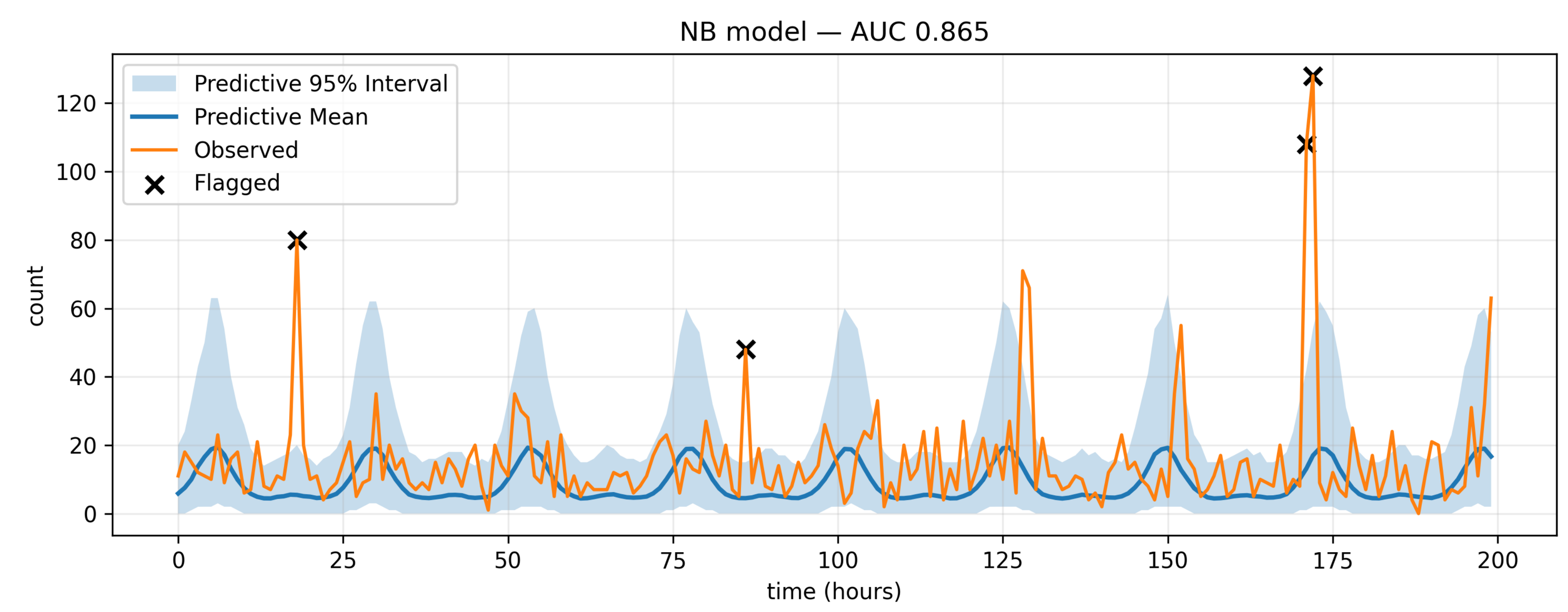


Figure 1. NB model — Posterior Predictive (AUC = 0.875)

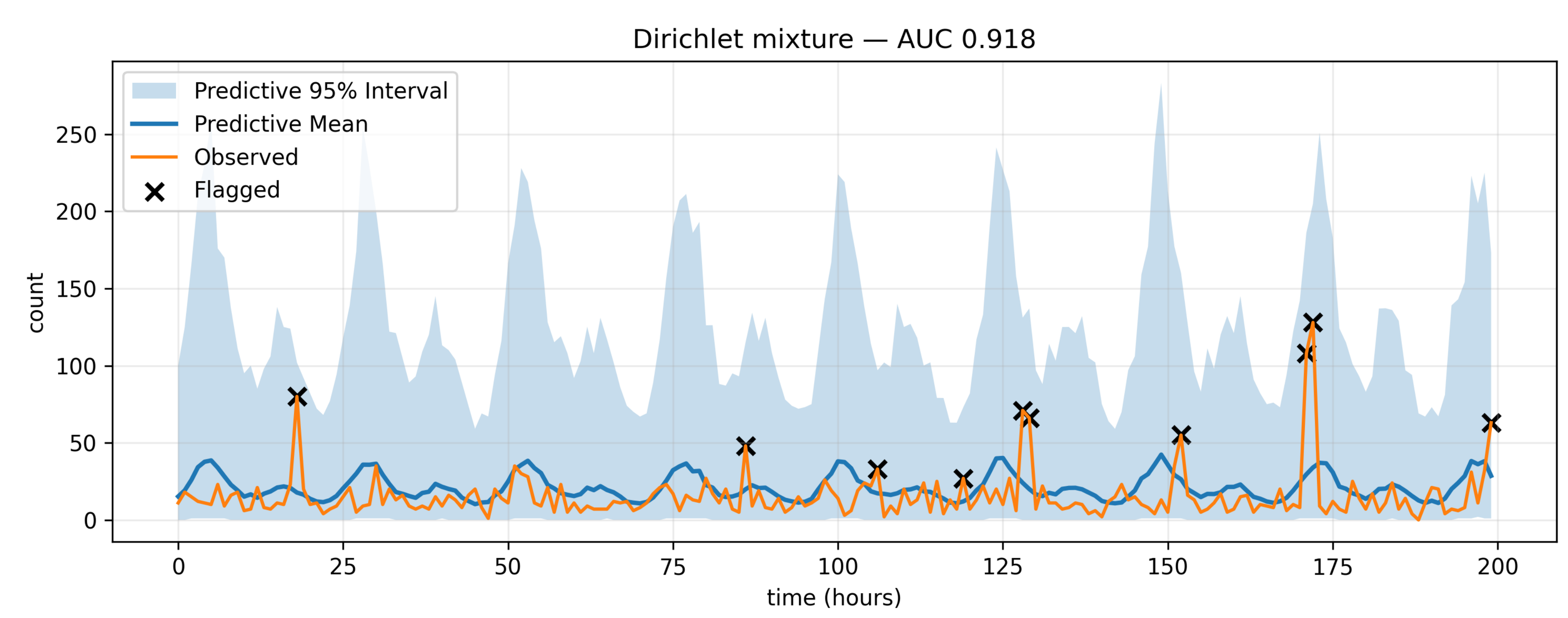


Figure 2. Dirichlet Mixture — Posterior Predictive (AUC = 0.912)

The two posterior predictive plots compare the single NB model and the Dirichlet–Mixture model, highlighting the improved ability of the mixture to capture rare and extreme spikes in IoT traffic.

Limitations & Next Steps

- Extend from single-device to multi-host detection with shared priors for coordinated alerts.
- Capture multi-hour intrusion patterns using temporal or HMM-based modeling.
- Incorporate network context and enable online Bayesian updates for real-time monitoring.

Summary

This work introduces a Dirichlet–Mixture Negative Binomial model for detecting rare spikes in seasonal IoT traffic. The proposed model explicitly separates *normal* and *spike* regimes through Dirichlet mixing, improving AUC, and interpretability compared to a standard NB approach.

References

[1] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 3 edition, 2013.

[2] Radford M Neal. Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics*, 9(2):249–265, 2000.

[3] José A. Perusquía, Jim E. Griffin, and Cristiano Villa. Bayesian models applied to cyber security anomaly detection problems. *arXiv preprint arXiv:2003.10360*, 2020. doi: 10.48550/arXiv.2003.10360. URL `https://arxiv.org/abs/2003.10360`.

[4] PyMC Contributors. Pymc: Probabilistic programming in python. `https://www.pymc.io`, 2024. Accessed: 2025-09-15.