

Supplementary Material for BARO: Robust Root Cause Analysis for Microservices via Multivariate Online Bayesian Change Point Detection

Figure S1 presents the experimental results of our proposed method, BARO, and the baseline root cause analysis methods: N-Sigma, ϵ -Diagnosis, CIRCA, and RCD on the Sock Shop dataset to evaluate their sensitivity w.r.t. the anomaly detection time \hat{t}_A . This figure is used in Section 4.8.1 to answer part 1 of RQ4 on the sensitivity of the anomaly detection time \hat{t}_A . Let us denote t_{inject} as the failure injection time. We formulate the anomaly detection time \hat{t}_A as $\hat{t}_A = t_{\text{inject}} + t_{\text{bias}}$ where t_{bias} ranges from -40 to 40. We then evaluate the performance of the methods with different values of \hat{t}_A within this range.

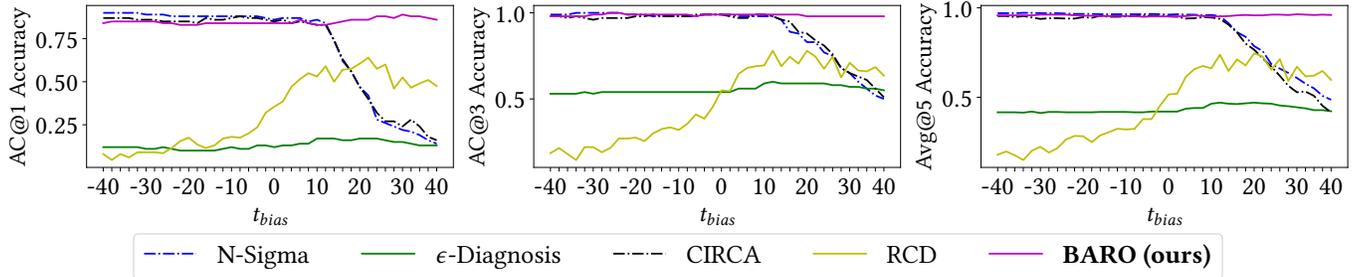


Figure S1. The performance of N-Sigma, ϵ -Diagnosis, CIRCA, RCD, and BARO w.r.t. different values of t_{bias} on the Sock Shop dataset. The figure presents the AC@1, AC@3, and Avg@5 scores from left to right.

Figure S2 reports the experimental results of three root cause analysis methods: CIRCA, RCD, and ϵ -Diagnosis on the Sock Shop dataset to evaluate their sensitivity w.r.t. their hyperparameters (significance threshold α in CIRCA and ϵ -Diagnosis, and chunk size γ in RCD). This figure is used in Section 4.8.2 to answer part 2 of RQ4 on the hyperparameter sensitivity analysis of RCA methods.

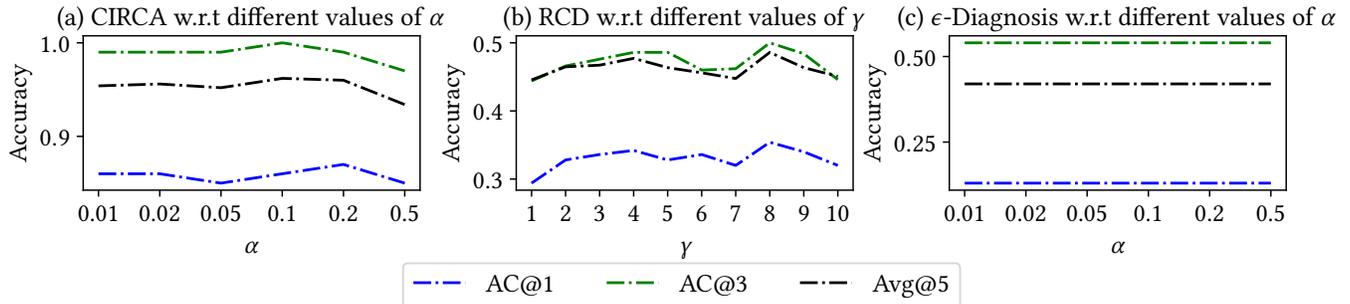


Figure S2. The performance of CIRCA (a), RCD (b), and ϵ -Diagnosis (c) w.r.t. their different parameter values on the Sock Shop dataset. The figure presents their AC@1, AC@3, and Avg@5 scores.