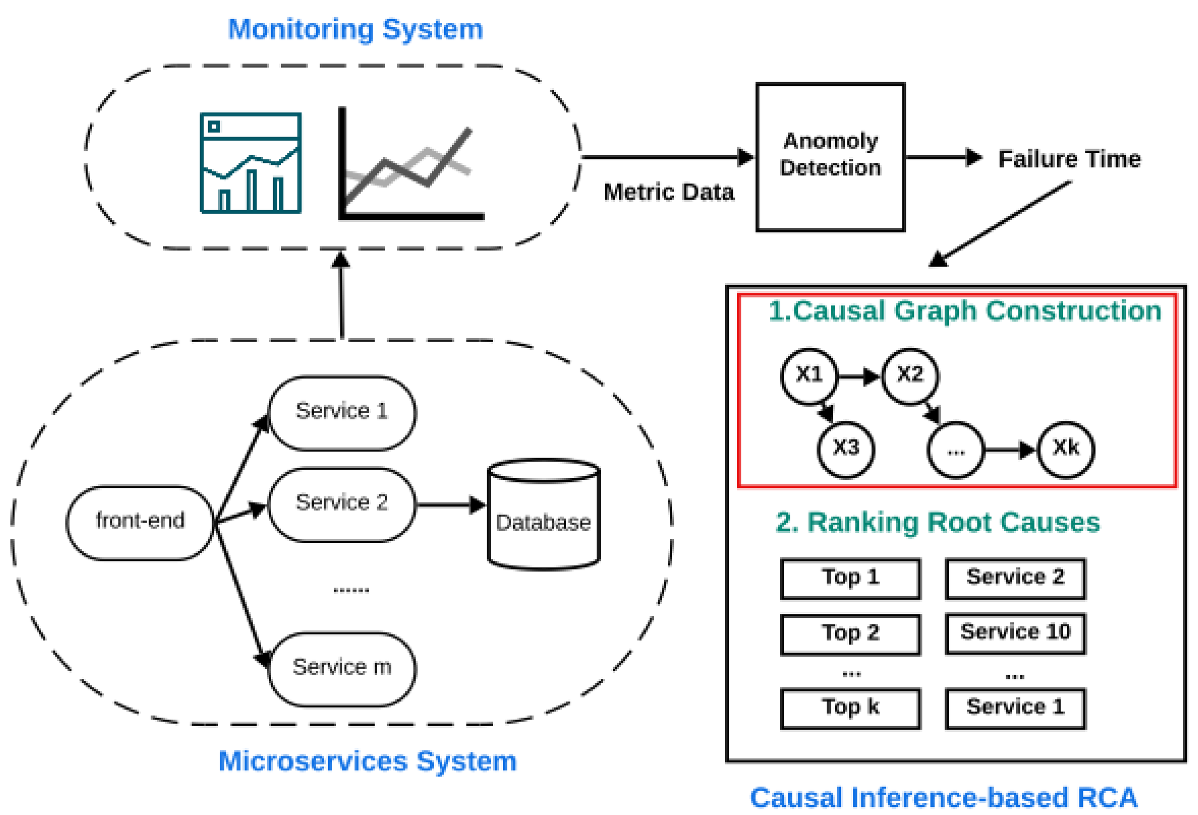
Root Cause Analysis of Failures in Microservices via Bayesian Root Cause Discovery

Kenneth Lee, Zihan Zhou, Murat Kocaoglu

Purdue University, Johns Hopkins University

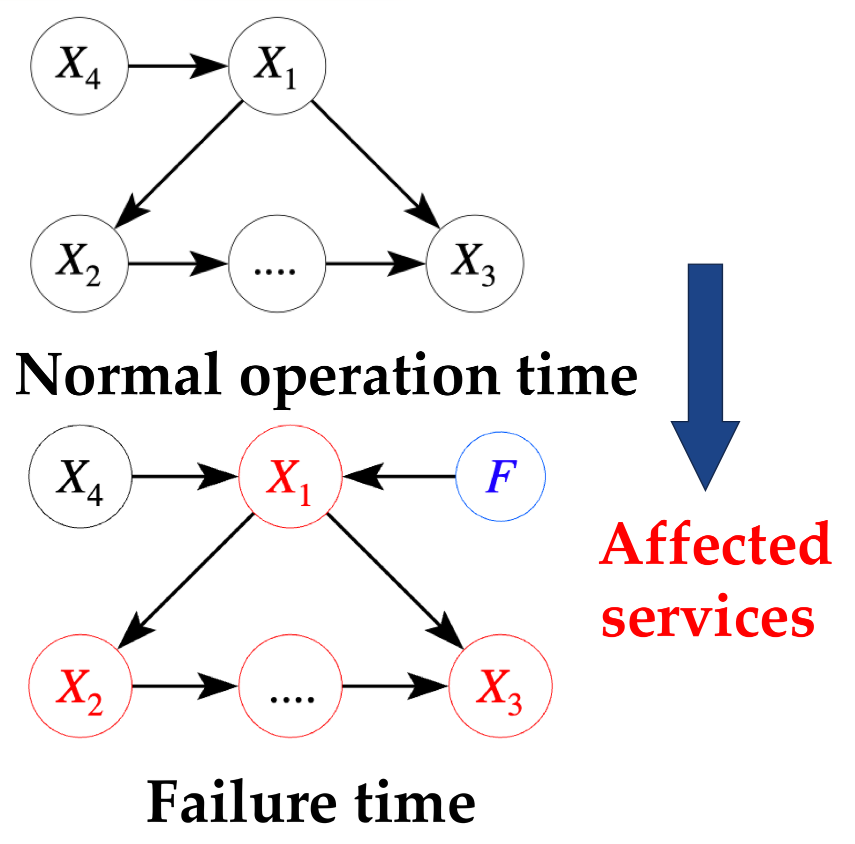
What is Root Cause Analysis?



Issues:
- ✕ Sample Inefficiency
- ✕ Ranking based on arbitrary functions.

Solution:
- ✓ Learn a partial causal structure **before** failure.
- ✓ Rank by posterior score function that depends on failures

Modeling Failures as Interventions



Bayesian Inference of the Root Causes

$$P(R\,|\,Data) = \sum_{C} P(R\,|\,C, Data)P(C\,|\,Data)$$

$$P(R\,|\,C, Data) = \frac{P(Data\,|\,R, C)P(R\,|\,C)}{\sum_{R'} P(Data\,|\,R', C)P(R'\,|\,C)}$$

$$P(C\,|\,Data) \propto \frac{P(Data\,|\,C)P(C)}{q(C)}$$

Evaluate the likelihood by sampling a DAG from C

Uniform over all sampled CPDAGs

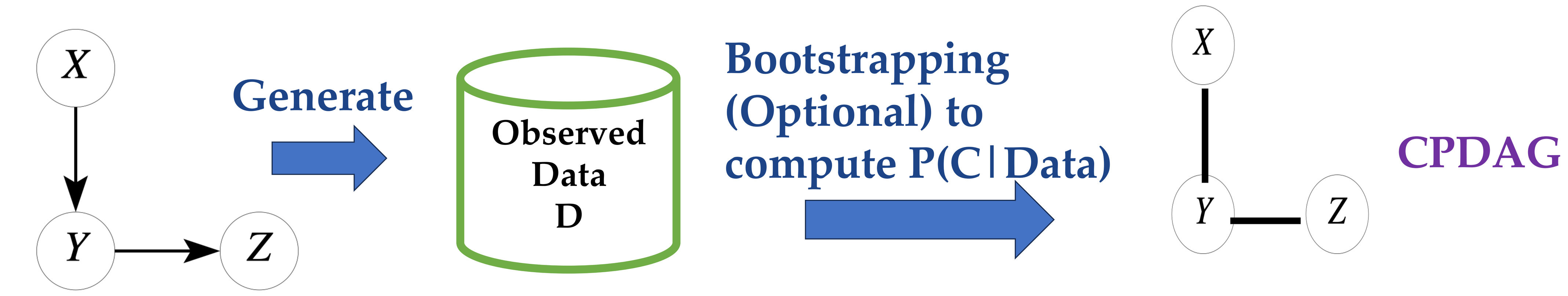
Empirical probability from bootstrapping

$$P(Data\,|\,R, C) = \sum_{G \in C} P(Data\,|\,G, R)P(G\,|\,C, R)$$

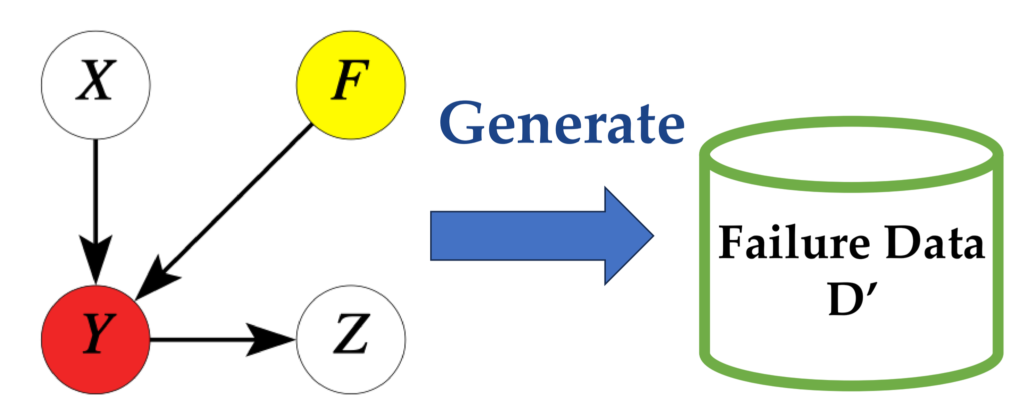
Proposed Algorithm (BRCD)

Before Failure

Step 1: Observational Causal Discovery

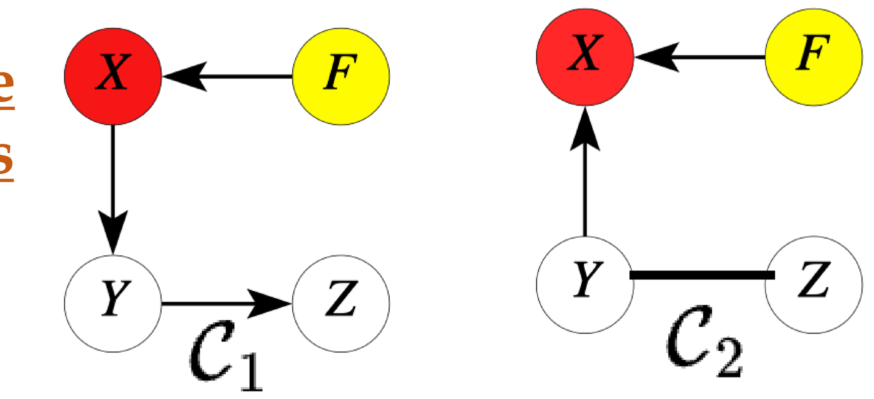


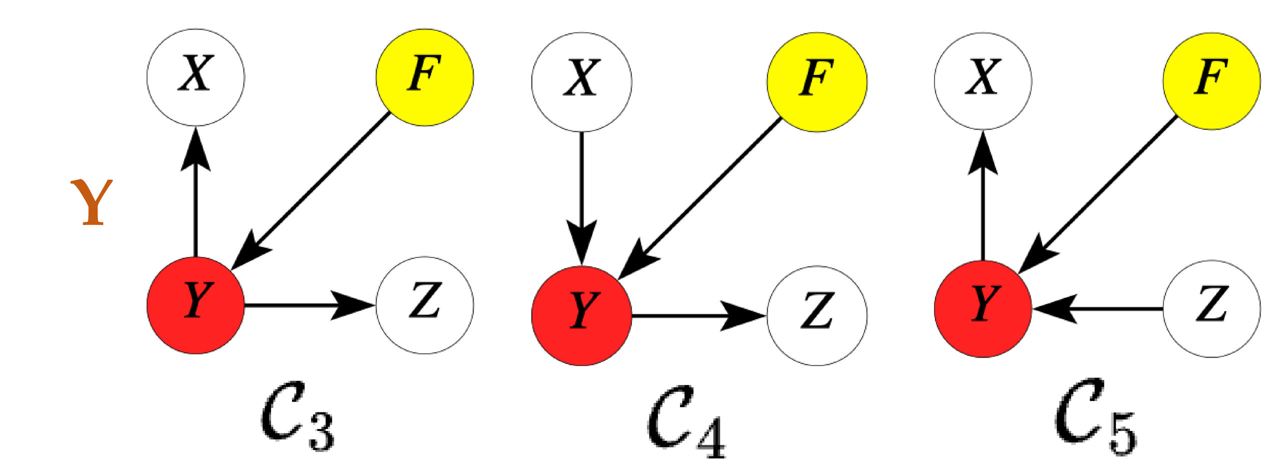
After Failure

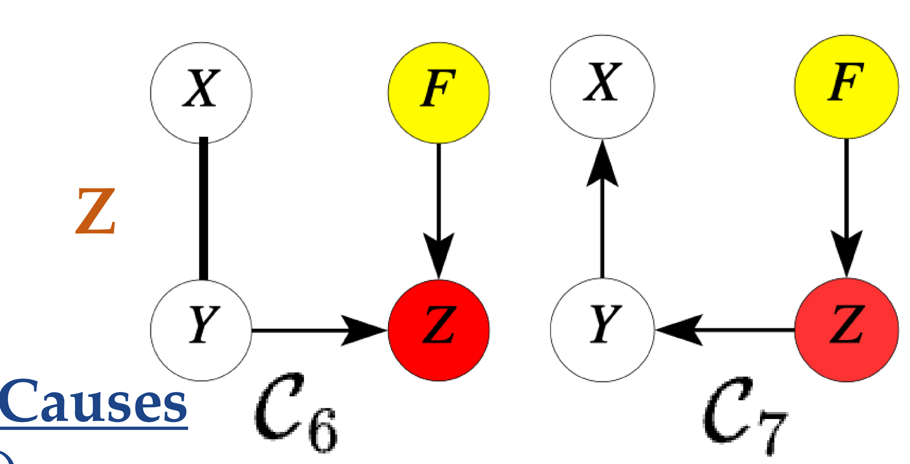


Step 2. Enumerate I-CPDAGs $\mathcal{C}_i$

Root Cause Candidates







Sampling DAG and counting the size of MEC only takes polynomial time [2]

Step 3: Compute Likelihood by Sampling DAGs

For a candidate X

$$\begin{aligned} &P(D|R = X) \\ &= P(Z|Y)P(Y|X)P(X|F)P(F)(1/9) \\ &+ P(Z|Y)P(Y)P(X|Y,F)P(F)(2/9) \end{aligned}$$

Factorize according the augmented DAGs sampled from $\mathcal{C}_1$ and $\mathcal{C}_2$

Step 4: Rank Root Causes by Posterior $P(R|D)$

Robustness to Finite Sample Approximation

Lemma (identifiability). Under modularity, positivity, for almost all parameter values, any two district sets R≠R′, if (G*,R*) is the ground truth, then

$$\Delta_{\min} := \inf_{(\mathcal{G},\mathbf{R})\neq(\mathcal{G}^\star,\mathbf{R}^\star)} \mathbb{E}_{p^\star}\left[\log \frac{p(\mathbf{X}|\mathcal{G}^\star,\mathbf{R}^\star)}{p(\mathbf{X}|\mathcal{G},\mathbf{R})}\right] > 0.$$

Assumptions
- A1 (ε-close plug in): With probability at least $1-\eta$, $\eta \in (0,1)$, n is the sample size, $d_{TV}(p', p^*) \le \varepsilon$, where $\varepsilon \to 0$ as $n \to \infty$
- A2 (Bounded log-likelihood ratios). There exists B $\le \infty$ such that for all (G, R)

$$\left|\log \frac{p(\mathbf{X}|\mathcal{G}^\star,\mathbf{R}^\star)}{p(\mathbf{X}|\mathcal{G},\mathbf{R})}\right| \le B.$$

Theorem (Posterior Consistency). Let R* be a set of root causes. Under A1-A2 and causal sufficiency, we have

$$p(\mathcal{G}^\star, \mathbf{R}^\star \mid \mathcal{D}) \xrightarrow[n\to\infty]{p^\star} 1.$$

Corollary. For any candidate target set R, BRCD enumerates all and only the I-CPDAG compatible with (C(G*), R) by ranging over all possible orientations of the edge cut E[R, V\R] and closing under Meek rules. Hence, BRCD covers all I-MEC consistent with (C(G*), R)

Theorem (Finite Sample Bound with ε robustness). For any $\delta \in (0,1)$, let M be the number of wrong pairs (G, R) and let

$$\Delta^{eff}_{min}(n) := \Delta_{\min} - 2B\varepsilon, t_n := B\sqrt{\frac{2\ln(2M/\delta)}{n}}$$

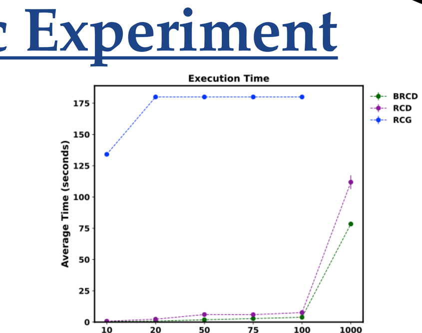
, with probability at least $1-\delta-\eta$,

$$p(\mathcal{G}^\star, \mathbf{R}^\star \mid \mathcal{D}) \ge 1 - M\exp\left\{-n\left(\Delta^{eff}_{min}(n) - t_n\right)\right\} \max_{(\mathcal{G},\mathbf{R})\neq(\mathcal{G}^\star,\mathbf{R}^\star)} \frac{p(\mathcal{G},\mathbf{R})}{p(\mathcal{G}^\star,\mathbf{R}^\star)}.$$

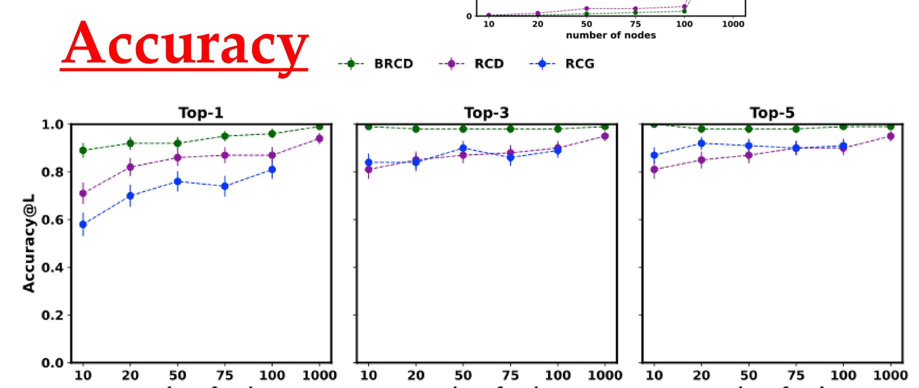
Lemma. Let G be a DAG and R be a set of root cause candidates. Consider (G, R) as a joint event. With a uniform prior on (G, R), ranking the posteriors of (G, R) is the same as ranking I-CPDAGs.

Synthetic Experiment

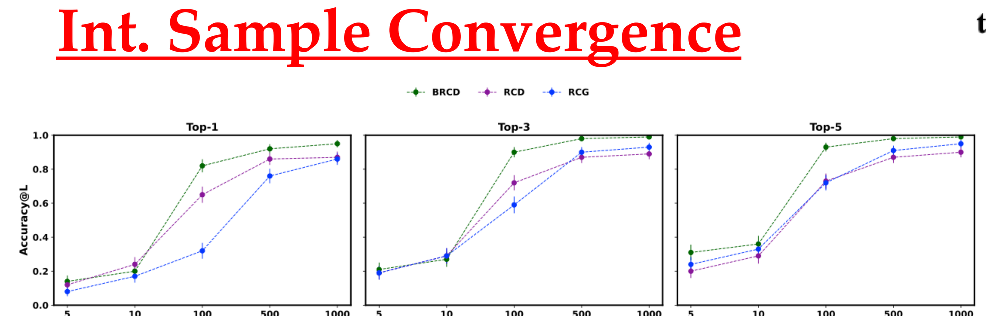
Scalability



Accuracy



Int. Sample Convergence



Experimental Results

Real-world Experiment with Petshop Data

		RCD	RCG	BARO	BRCD	BRCD-C	BRCD-M	BRCD-B10	BRCD-B100	SO	ST	Cholesky	IDI	ShapleyIQ	MicroDig	SimpleRCA
	high_traffic	0.00	0.00	0.15	0.27	**0.35**	0.08	0.31	**0.31**	0.00	0.00	0.00	0.12	0.08	0.00	0.08
	low_traffic	0.31	0.15	0.27	**0.46**	0.42	0.35	0.42	**0.46**	0.04	0.00	0.00	0.12	0.15	0.00	0.15
top-1	temporal_traffic1	0.25	0.12	0.25	0.38	**0.63**	**0.63**	0.25	0.50	0.00	0.00	0.00	0.25	0.12	0.00	0.25
	temporal_traffic2	0.62	0.25	0.25	0.50	**0.88**	**0.88**	0.62	0.62	0.00	0.00	0.00	0.25	0.25	0.00	0.25
	Average	0.30	0.13	0.23	0.40	**0.57**	0.48	0.40	0.47	0.01	0.00	0.00	0.19	0.15	0.00	0.18
	high_traffic	0.00	0.00	0.23	0.35	**0.38**	0.23	0.31	0.35	0.00	0.00	0.00	0.23	0.23	0.42	0.19
	low_traffic	0.38	0.19	0.42	0.65	0.69	0.58	**0.77**	0.73	0.04	0.04	0.00	0.23	0.38	0.62	0.31
top-3	temporal_traffic1	**0.75**	0.25	0.38	0.62	0.75	0.75	0.62	**0.75**	0.12	0.00	0.00	0.38	0.38	0.62	0.38
	temporal_traffic2	0.75	0.25	0.38	0.62	**0.88**	**0.88**	**0.88**	**0.88**	0.00	0.00	0.00	0.50	0.50	0.62	0.38
	Average	0.47	0.17	0.35	0.56	**0.68**	0.61	0.65	**0.68**	0.04	0.01	0.00	0.34	0.37	0.57	0.32
	high_traffic	0.00	0.23	0.31	0.38	**0.42**	0.38	0.35	**0.42**	0.00	0.00	0.00	0.35	0.31	0.62	0.35
	low_traffic	0.42	0.31	0.46	0.73	0.73	0.58	0.81	**0.85**	0.08	0.23	0.04	0.35	0.46	0.65	0.46
top-5	temporal_traffic1	**0.75**	0.38	0.38	0.62	**0.75**	**0.75**	0.62	**0.75**	0.25	0.00	0.00	0.38	0.38	0.62	0.38
	temporal_traffic2	0.75	0.38	0.50	0.75	**0.88**	**0.88**	**0.88**	**0.88**	0.00	0.00	0.12	0.62	0.62	0.62	0.38
	Average	0.48	0.33	0.41	0.62	0.69	0.65	0.67	**0.73**	0.08	0.06	0.04	0.43	0.44	0.63	0.39

Reference

[1] Ikram, Azam, et al. "Root cause analysis of failures in microservices through causal discovery." Advances in Neural Information Processing Systems 35 (2022): 31158-31170.

[2] Wienöbst, M., Bannach, M., and Li´skiewicz, M. Polynomial-time algorithms for counting and sampling markov equivalent dags with applications. Journal of Machine Learning Research, 24(213):1–45, 2023.

Paper **Code**