Journey to the Centre of Cluster: Harnessing Interior Nodes for A/B Testing under Network Interference

Qianyi Chen (THU), Anpeng Wu (ZJU), Bo Li* (THU), Lu Deng, Yong Wang (Tencent)

Scan for paper:

A/B Testing under Network Interference

A unit's potential outcome is not only decided by its own assigned treatment, but also treatments assigned to its neighbors on network.

With classical neighborhood interference assumption (in violation of SUTVA), we have:

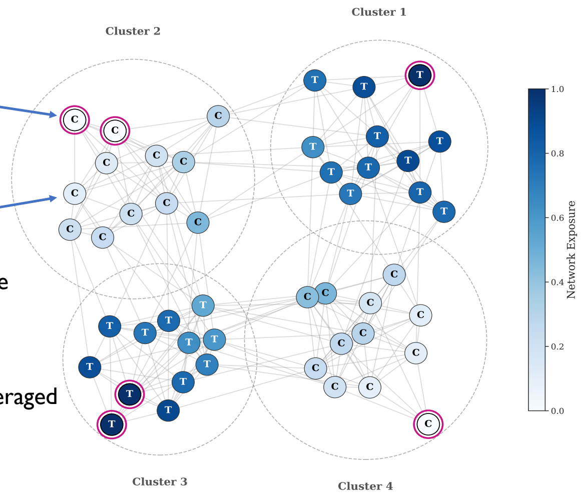
$$Y_i = Y_i(z_i; \{z_j : A_{ij} = 1\})$$

Leveraging Interior Units: A Tale of Semi-Supervision in Network Experiments

High-quality labeled sample: interior nodes, whose exposure levels always equal 0/1 and can be viewed as ground-truth in estimating Global ATE.

Low-quality pseudo-labeled sample: boundary nodes, the majority of whole population.

Though being biased samples due to network interference, they can be leveraged to reduce the selection bias of cluster interior, through PPI-type adjustment.



Problem of Existing Estimators

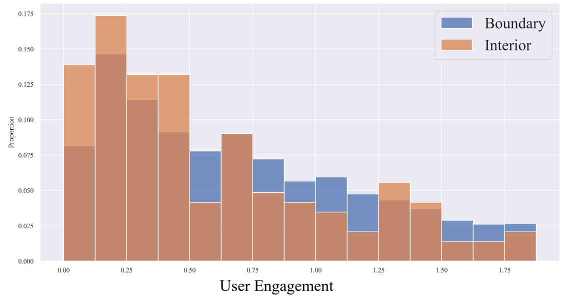
- Overdependency on interior nodes
- High-variance for statistical estimators, and high-bias for regression-based estimator

Mean-in-Interior (MII) Estimator: Difference-in-Means for Interior Nodes

$$\hat{\tau}_{MII} = \frac{\sum_{i \in \mathrm{Int}} z_i Y_i}{\sum_{j \in \mathrm{Int}} z_j} - \frac{\sum_{i \in \mathrm{Int}} (1 - z_i) Y_i}{\sum_{j \in \mathrm{Int}} (1 - z_j)}$$

- Reduces variance with moderate weights (in comparison to inverse propensity weight, exponentially scale w.r.t. a unit's degree).
- Still incurs selection bias due to discrepancy between interior and boundary units.

Empirical evidence of discrepancy



Augmented MII Estimator: Leveraging Boundary Nodes for Debiasing Adjustment

$$\hat{\tau}_{AMII} = \hat{\tau}_{MII} + \left(\frac{1}{n} \sum_{j \in [n]} f(\mathbf{1}, X, A)_j - \frac{1}{s_1} \sum_{i \in \mathrm{Int}} z_i f(\mathbf{1}, X, A)_i \right) - \left(\frac{1}{n} \sum_{j \in [n]} f(\mathbf{0}, X, A)_j - \frac{1}{s_0} \sum_{i \in \mathrm{Int}} (1 - z_i) f(\mathbf{0}, X, A)_i \right)$$

- Train a counterfactual predictor f for removing the selection bias of MII estimator.
- Intuition: regression component for capturing covariates; MII for capturing network exposure at two terminals (global 0/1). Debiasing using predictions.

Experiments performance: Bias, Std, MSE