Contamination-Aware Experimentation on Networks

Mine Su Erturk (mserturk@stanford.edu), Eray Turkel (eturkel@stanford.edu)

Motivation

We study a setting where a decision maker is conducting experiments in a network environment. We assume the existence of multiple analysts conducting experiments on the same network, as is the case in many online platforms.
An experiment creates negative externalities on other ongoing experiments by contaminating their results. We analyze an experimenter's decision making problem in this setting, where the goal is learning an optimal treatment regime over the network while limiting the contamination on other experimenters. We provide theoretical regret bounds and study the performance of our suggested policy through simulations.

Exposure of node 'i' at time t

$$e_{it} = \frac{\sum_{j \leqslant N} \mathbf{1}(j \in nhbd(i)) w_{jt}}{\sum_{j \leqslant N} \mathbf{1}(j \in nhbd(i))}$$

$w_{j,t}$= Treatment status of node j at time t

Contamination cost at time t

$$c_t = c \sum_{i \leqslant N} \mathbf{1} \left(e_{it} > 0 \text{ and } e_{is} = 0, \forall s < t \right)$$

Contamination: Affecting the results of Other Experimenters

The DM's goal is choosing an optimal treatment regime, while staying below a contamination budget in order to limit negative externalities on other experimenters. We assume that the cost is proportional to the number of nodes exposed to the treatment, and once a node is exposed, future exposures of that unit do not increase the cost any further.

Outcome for node 'i' at time t

$$Y_{it}(w_{it}, e_{it}) = \beta^T x_i + \Gamma^T x_i e_{it} + \epsilon_i$$

Choosing whom to treat: LP approach

Given a history of experimentation and observed outcomes, we can characterize the DM's problem of choosing a treatment deployment (choosing which **D** nodes to treat) over the network as an integer program, and study its relaxation as an LP. We interpret the resulting optimal policy as a probabilistic treatment assignment rule.
The DM will have an estimate of the treatment response at the end of experimentation, which we call $\hat{\Gamma}$. This estimate will be used to maximize an empirical value function.

$$Max_{\hat{\pi}} : \sum_{i \leqslant N} \hat{\Gamma}^T x_i e_i(\hat{\pi}) \quad \text{s.t. } 1^T \hat{\pi} \leqslant D$$

We write this as an LP using the adjacency matrix of the network **A**, and the matrix of observed covariates, **X**.

$$Max_{\hat{\pi}} : \left(\hat{\Gamma}^T X Diag(A\mathbf{1}_N)^{-1} A \right) \hat{\pi}, \text{ subject to: } 1^T \hat{\pi} \leqslant D, \forall i, \hat{\pi}_i \in [0,1]$$

We analyze regret **R,** as a function of the contamination budget $\bar{C}$**,** against an oracle who knows the true value of the parameter Γ. We use perturbation analysis of the linear program to show the following result.

Theorem 1 *Fix $\tau \geqslant 0$. Then, regret is bounded above by K with the following probability:*

$$\mathbb{P}(R \leqslant K) \geqslant 1 - \exp\left(-\frac{1}{2}(\tau - 1)^2 \right),$$

where

$$K = \frac{\delta \max\{\sqrt{n + D^2}, ||\Gamma||\sigma_1(XX^T)\}}{\sqrt{n+1}} \cdot \frac{\sqrt{n+D^2}}{\sqrt{n+1} - \delta},$$

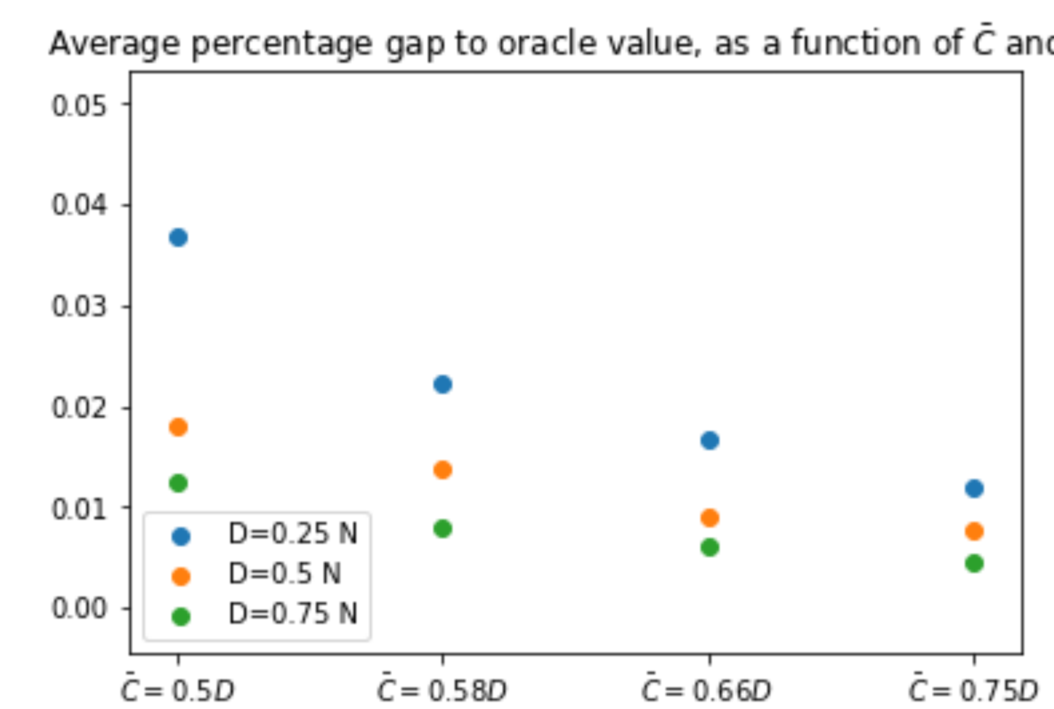
and $\delta = \tau \sigma_\epsilon \sum_{j=1}^k \sigma_j((X_s X_s^T)^{-1})$.

Regret depends on the first principle component of the covariate matrix of the entire network $\sigma_1(XX^T)$, and decreases with information gained in the experimentation period through the term $\sum_{j=1}^{k} \sigma_j (X_s X_s^T)^{-1}$.

Choosing where to experiment: Knapsack Approach

Our proposed algorithm chooses the most valuable nodes to experiment on greedily at every period, taking into account the cost of contaminating other nodes. We measure the information gain from experimenting on a given node by the change in the trace of the matrix $(X_s X_s^T)^{-1}$ after adding all the unexposed neighbors of the target node to the set of exposed nodes. We calculate the cost of experimenting on a node by calculating the number of new exposures resulting from treating the target node. We get the value of experimenting on a node by taking the ratio of the information gain to the cost, and determine the path of experimentation by sequentially choosing the nodes with the highest value until we reach our contamination budget, $\bar{C}$.

Simulation Experiments on Random Graphs



Percentage gap to oracle value, for a set of different contamination budgets in simulation studies on Erdos-Renyi random graphs with 1000 nodes, and independent link formation probability p=0.1 between every node. A random 10-dimensional unknown parameter vector Γ is drawn from a Gaussian distribution for every simulation. The covariate matrix for the nodes are also drawn from a Gaussian distribution, assuming no correlation between nodes. 1000 simulations were run for every parameter combination.

Future Work

We will focus on understanding how the network topology impacts the tradeoff between contamination and learning. We will also analyze how the performance of the DM is affected by homophily in the network, if the observed covariates of neighboring nodes were correlated. Studying the strategic interaction between multiple experimenters sharing the same subject pool with and without a central planner remains as an interesting open question.