

Contamination-Aware Experiments on Networks

Mine Su Erturk
Graduate School of Business
Stanford University
mserturk@stanford.edu

Eray Turkel
Graduate School of Business
Stanford University
eturkel@stanford.edu

1 Introduction

In this project, we study the problem of designing an experimentation platform, while taking into account possible contamination between experimenters. In reality, many teams operate on the same subject pool, each of them testing their own treatments and possibly interested in different outcome metrics. However, the objective of the platform is finding the best combination of these treatments, rather than making independent ship/no-ship decisions on each experiment individually. This creates a tension between the incentives of the platform and individual experimenters. While each group might make statistically sound decisions on whether to ship their experiment in isolation, the best combination of all the treatments, which is what the platform is interested in, might be missed.

The goal is to follow a mechanism-design based approach to build an experimentation platform for multiple teams working on the same subjects. Through our analysis, we plan to develop insights to guide the design of experimentation platforms that will aim to internalize the externalities which experimenters impose on each other, hence aligning the incentives of each individual experimenter and the platform as a whole.

–motivate the network

When is the knowledge of the underlying interaction graph between experiments useful? Are there regimes in which graph-agnostic policies perform as well as their graph-aware counterparts?

–Further, studying the interaction between multiple experimenters over a network with and without a central scheduler remains as an interesting open question for our future work.

2 Introduction-old

We study the problem of a decision maker (DM) conducting experiments over time, on a network environment. An ‘experiment’ in the context of our model is treating a node in the network. Each node responds to the treatment differently, depending on their (observable) characteristics. We also assume that treating a node creates spillovers on neighboring nodes. The DM’s goal is learning the optimal treatment allocation over the network by sequentially conducting local experiments on different nodes.

We assume that there are multiple analysts conducting experiments on the network, as is the case in many online platforms. The DM’s experiment creates negative externalities on other ongoing experiments, or the deployed treatments from previously run experiments. We model this as a cost for contaminating the treatment regime of other experimenters. Therefore, the DM has to strike a balance between learning the treatment response function efficiently, and staying within acceptable levels of contamination. We analyze the DM’s problem as a sequential decision making problem, which puts our framework closer to the literature on online learning and bandits, instead of using a randomization based inference setting, as is common in the network experimentation literature.

3 Model-old

There are N units connected on an undirected, unweighted graph. Denote each unit on the graph by $i \in \{1, \dots, N\}$. The units have observable characteristics, $\{x_1, \dots, x_N\}$ with $x_i \in \mathbb{R}^k$. The DM’s

objective depends on the outcomes of the units, denoted $\{Y_1, \dots, Y_N\}$ with $Y \in \mathbb{R}$, which respond to a treatment regime that will be determined. The decision maker can run experiments sequentially, at time periods $t \in \{1, \dots, T\}$. Let w_{it} denote the treatment assignment for node i during the period of experimentation t , with $w_{it} \in \{0, 1\}$, and denote the vector of treatment assignments at time t by $w_t = [w_{1t}, \dots, w_{Nt}] \in \{0, 1\}^N$. We assume that the experimentation is local, in the sense that every period, only one node can be treated. Thus for all time periods t , $\sum_{i \leq N} w_{it} = 1$.

Each node's outcome depends on their characteristics, the treatment they receive and the exposure to the treatment through their neighbors. Define the exposure for node i at period t as e_{it} , which is assumed to be:

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

Y is assumed to be a linear function of the unit characteristics and the exposure to the treatment:

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

where $\beta, \Gamma \in \mathbb{R}^k$, and for all nodes i , $\epsilon_i \sim N(0, \sigma_\epsilon^2)$.

The DM's goal is to choose a treatment assignment policy over the network based on the value of Γ . Denote the final treatment assignment for node i by w_i , dropping the time subscript. We denote a policy by $\pi(\Gamma)$, where $\pi(\Gamma) = [w_1, \dots, w_N] \in \{0, 1\}^N$. We assume that there is an upper limit on the number of nodes that can be treated in the final deployment, which we call D , i.e., $\sum_{i \leq N} w_i \leq D$.

At every period t , the DM's experimentation creates a cost, because the experiment interacts with a previous experimenter's treatment allocation, contaminating their results. We assume that the cost is proportional to the number of units exposed to the treatment, and once a unit is exposed, future exposures of that unit do not increase the cost any further. Define the total cost of experimentation at period t by:

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

Throughout, we will let $c = 1$ for simplicity. Hence, the total cost incurred through experimentation by the DM can be written as $C = \sum_{t \leq T} c_t$. We assume that the DM has to stay below a 'contamination budget' of at most \bar{C} during experimentation. Then, as a function of the contamination budget \bar{C} , we define the value of a policy as follows:

$$V(\pi, \bar{C}) = \sum_{i \leq N} \Gamma^T x_i e_i(\pi),$$

where $e_i(\pi)$ denotes the exposure function determined at the last period for final deployment, induced by the treatment allocation under policy π , decided after an experimentation period with the contamination budget \bar{C} . The DM does not observe the true vector Γ but has an imperfect estimate of it learned through sequential experimentation, which we will call $\hat{\Gamma}$. The DM maximizes the empirical value function using its estimate at the end of the experimentation period:

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

Let $X_{K \times N}$ denote the matrix of covariates for all the nodes on the graph, and let X_s denote the analogous matrix only for nodes that were exposed to the treatment during the experimentation period. Let $\hat{\Gamma}_{1 \times K}$ be the vector of estimated coefficients.

Define the vector of exposure values at time t as $\bar{e}_t = [\mathbf{1}(e_{1t} > 0), \dots, \mathbf{1}(e_{Nt} > 0)]_{1 \times N}$ and let c_t denote the cost vector at time t . We can recursively write: $c_0 = \mathbf{0}_{1 \times N}$, $c_1 = \bar{e}_1$, and generally, $c_n = (c_{n-1} + \bar{e}_n) - (\text{Diag}(c_{n-1})\bar{e}_n)$. $\text{Diag}(c_n)$ denotes the diagonal matrix with the entries of c_n on its diagonal entries. This recursive formulation captures the notion that the cost of second and future exposures to the experiment are zero.

We can also define the final exposure vector e in terms of the adjacency matrix of the network, $A_{N \times N}$ (with self-edges), and the final treatment allocation vector $\pi \in \{0, 1\}^N$. Representing e as a

vector, we have: $Diag(A\mathbf{1}_N)^{-1}A\pi = e$, where $\mathbf{1}_N$ is an $(N \times 1)$ vector of 1's. Therefore, we can rewrite the maximization problem (1) as a convenient linear program, where we relax π_i and allow it to take values in $[0, 1]$, interpreting the resulting policy as a probabilistic treatment assignment rule:

$$Max_{\hat{\pi}} : \left(\hat{\Gamma}^T X Diag(A\mathbf{1}_N)^{-1} A \right) \hat{\pi}, \text{ subject to: } \mathbf{1}^T \hat{\pi} \leq D, \forall i, \hat{\pi}_i \in [0, 1]. \quad (2)$$