

Contamination-Aware Experiments on Networks

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

1.1 Overview of the Model

Suppose the platform has multiple metrics they care about such as user retention, user growth, or ad revenue. We assume that the objective of the platform can be written as a linear combination of these metrics. The platform delegates experimentation to multiple teams with each team focusing on developing interventions that will affect a subset of these metrics. For example, one team might focus on increasing the average time spent on the platform whereas another team might design interventions to limit the spread of harmful content on the platform.

The teams obtain positive payoff if and only if their proposed interventions are shipped (or implemented). Moreover, there is a possibility that a team's treatment interacts with the focal metrics of another team. We represent this interaction of interventions through a potentially unknown graph $G = (V, E)$. We assume that there is a node for each metric-team pair. In other words, we create "virtual copies" of each metric for every team on the graph. There exists an edge between $(metric_i, team_j)$ and $(metric_k, team_l)$ if and only if the proposed interventions of team j and l interfere with each other. For simplicity, we will begin by assuming that a team's intervention can only interfere with another team through the same metric, i.e., edges can only exist between $(metric_i, team_j)$ and $(metric_i, team_l)$.

Shipment decisions are made according to a linear decision rule, $\mathbb{I}(\beta Y \geq 0)$ where Y_i is the change in metric i and β is the coefficient vector over metrics. In a decentralized setting, each team is allowed to choose their own β vector which means that they can focus on different metrics to maximize their probability of getting shipped. Our goal as a platform is to design a system to restrict the choices of β vectors such that the individual decisions of the teams are aligned with the overall objective of the platform. This implies that the optimal choice of the β vectors might depend on the underlying interaction graph. Our main research questions are:

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