COMS 6998-13: Analysis of Social Networks
Liushiya Chen (lc3501)

Research Proposal:
Minimizing Spillover in Experiments from a Network Theory Perspective

ABSTRACT

In an increasingly connected world, social science experiments involving a treatment and a control group can suffer from spillover when the treatment given to the treatment group indirectly affects the control group. Spillover is common in information-related treatments such as voter mobilization and advertising, because information can spread within social circles. There exist many theoretical frameworks to detect and estimate the extent of spillover, but I will view this problem from a social network perspective. Given a social network of experiment subjects, I propose formulating the treatment and control groups selection problem as an optimization problem to find the optimal selection of treatment and control subjects. Finding an efficient algorithm to solve this problem at scale will help social scientists design better experiments that estimate treatment effects more accurately.

INTRODUCTION

Experimental Design

Most experimental research in social science involves estimating the causal effects of a treatment. For example, a pharmacologist will want to estimate the effect of a new drug on subjects' blood pressure.

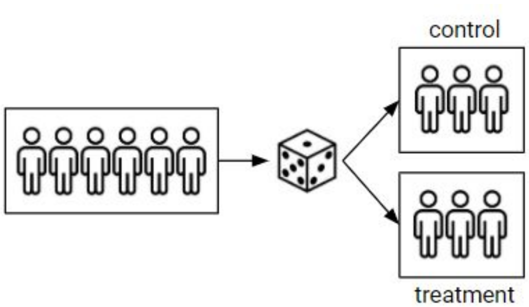

Figure 1: randomized control trial

The most widely used technique is randomized controlled trial (RCT), where subjects are randomly assigned into either the treatment or the control group. The treatment group will be given the treatment (e.g. a drug), while the control group will either be given a placebo or simply not given the treatment. The estimated causal effect of the treatment, also known as the average treatment effect (ATE) will be the difference in average outcome (e.g. blood pressure) between the control and treatment group:

$$A\hat{T}E = \bar{Y}_{treatment} - \bar{Y}_{control}$$

RCT is often referred to as the gold standard because RCT ensures that the group assignment is completely unbiased and the subjects in each group will likely be very similar when N is large.

Spillover

In some scenarios, the treatment can indirectly affect the control group, leading to misestimation of the true effect. For instance, a public health official wants to investigate the effect of anti-smoking ads on smoking. She randomly selects a group of subjects in a city to be shown the ads (treatment) and another to be not shown the ads (control). Treatment subjects then start talking to control subjects about the ads who are then more likely to quit smoking. This effect on the control group is the spillover.

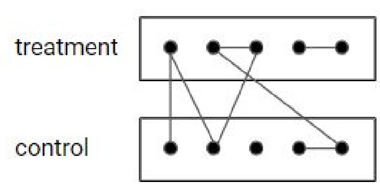

Figure 2: graphical representation of social connections (edges) between control and treatment subjects (nodes) that could lead to spillover

Suppose both seeing and hearing about the ads increase motivation to quit smoking, spillover will result in an underestimation of the average treatment effect. This is because the outcome in the control group is also improved indirectly as a result of the treatment, and the difference between the two groups is less drastic. The estimator for ATE below will be a biased estimator due to spillover that affects the outcome of a hypothetical "ideal control" where no spillover occurs.

$$A\hat{T}E = \bar{Y}_{treatment} - (\bar{Y}_{ideal\ control} + spillover)$$

EXISTING WORK

Multi-Level Experiments

Spillover is a well-known concept in statistics. Using a multi-level experiment is a common approach to reduce spillover. Sinclair et al. (2012) proposed a social network-inspired multi-level experiment that groups the subjects into precincts and then households before introducing randomization at both the precinct, household, and individual level. This provides a statistical test for the existence of spillover effect. Baird et al. (2014) proposed a multi-level experiment approach that splits the subjects into clusters and randomizes the intensity of the treatment within each cluster. This allows for a statistically consistent estimation of the spillover effect.

Interference Network

More recently, Aronow et al. (2020) considered an interference network of the subjects, which are students in the same school in his experiment. Each node in the interference network represents a subject, while each edge represents a possible channel through which spillover can happen between a pair of subjects. By studying the degrees of each node, Aronow et al. estimates the extent of spillover effects on each subject. This is the direction this proposal will expand upon.

OPTIMIZATION PROBLEM

While existing work provides a theoretical framework for detecting spillover and estimating the effects of spillover, little is done to answer the question of how to select the optimal treatment and control group from a pool of subjects. Here, I plan to build upon Aronow et al.'s (2020) idea by formulating the problem as a subset selection optimization problem.

Suppose we want to evaluate a binary treatment, i.e. subjects are either treated or not without any partial treatment. Suppose further that we know the interference network between subjects. Then, intuitively, an optimal assignment of treatment and control group should minimize the expected number of indirectly treated subjects in the control group.

More formally, the interference network is a weighted undirected graph G = {V, E, W} as described by Aronow et al. (2020), where V is the set of subjects and each edge E(i,j) is a channel through which spillover can occur between subjects i and j. Edge weight W(i,j) denotes the probability of a treatment spilling over from subject i to subject j over the duration of the experiment. We want to find treatment group $T \subset V$ and control group $C \subset V$, such that expected spillover is minimized.

$$\begin{aligned} T^*, C^* &= \arg\min_{T,C} \sum_{c \in C} P(c \text{ treated}) \\ &= \arg\min_{T,C} \sum_{c \in C} \left[1 - \prod_{n \in N_{G'}(c)} W(c,n)\right] \end{aligned}$$

Equation 3: we seek T* and C* that minimizes the expected number of treated control subjects. G' is the subgraph of G with nodes $T \cup C$, and $N_{G'}(c)$ denotes neighbor of subject c in G'.

We also assume there exists a target number of subjects M that we plan to treat. This gives us the following constraints:

$$T \cap C = \emptyset, \quad |T| + |C| \leq M$$

If the number of subjects is small, this problem can be solved by brute force. If the number of subjects is large, it is more feasible to investigate approximate algorithms with tight theoretical bounds.

PROBLEM OF HOMOPHILY

Unfortunately, another issue arises when we focus on minimizing spillover. McPherson et al. (2001) showed that homophily exists in social networks and there is more social interaction between similar subjects. This can create systematic differences between treatment and control subjects. This could interfere greatly with the average treatment effect estimation (Grossman & Mackenzie, 2005).

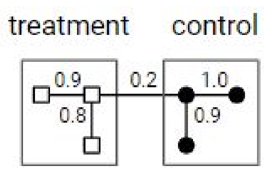

Figure 4: a potential treatment and control group selection with homophily in interference network

Suppose we have the interference network in Figure 4, where square nodes represent elderly subjects and round nodes represent young subjects. Minimizing the expected spillover will yield two homogenous groups in this case. If the treatment's effectiveness (e.g. voter mobilization) depends significantly on age, our estimates of the average treatment effect on the population will be skewed due to the algorithm's sole focus on reducing spillover.

PROBLEM REFORMULATION

In addition to spillover, we must also account for the homogeneity/similarity between control and treatment groups. I propose two approaches:

1. **Homogeneity as a constraint:** determine a minimum level of homogeneity between two groups we need then minimize spillover
2. **Homogeneity in objective function:** modify our objective function to be a weighted sum of expected spillover and homogeneity

There are many metrics to quantify between-group homogeneity, one of them could be comparing the histograms of each treatment group.

This gives a more complex optimization problem than before so an exact algorithm will likely not scale well for large social networks. Again, I propose exploring heuristics when using this algorithm for experiments with more than thousands of subjects and non-sparse interference networks.

APPLICATION OF RESULTS

Suppose we managed to develop a scalable algorithm that reduces spillover while maintaining a high level of similarity between control and treatment groups. The algorithm could help researchers design better experiments if these criteria are met:

1. There is a treatment with an unknown causal effect on a quantifiable outcome that researchers would like to estimate.
2. There is some information about the underlying interference network, i.e. there is some proxy to approximate how the treatment effect could spill over from one subject to another. An example would be Facebook friendship between subjects.
3. There is data on quantifiable factors (age, income, etc.) that could impact treatment effect and help to measure homogeneity.

Here are two examples of practical scenarios where this algorithm could be useful.

Scenario 1:

Social media platform's marketing team conducting A/B testing to determine if they should use a pop-up to promote their recently added feature. The treatment is whether a user is shown the pop-up and the outcome can be a KPI such as click rate. The interference network could be approximated by the number of interactions between each pair of users. Finally, the personal information of each user can be used to estimate homogeneity of groups of users. With this algorithm, the marketing team can partition users into treatment and control groups to estimate the effects of showing the pop-up.

Scenario 2:

A local government wants to promote a new SAT tutoring program and would like to test if mailing out pamphlets is effective at increasing enrollment rates. The spillover in this case will be awareness of the program spreading through word of mouth. An interference network could be built using the address of target parents and the schools attended by their children. Furthermore, demographic and academic information about both parents and their children could be used to determine homogeneity of subject groups. Using this algorithm, the government could decide which parents to mail the pamphlets to in order to best estimate the causal effect of sending pamphlets.

CONCLUSION

Social experiments can be messy and the classical assumptions in experiment design might not always hold. Given additional information about how the treatment effect could flow within a network of subjects, there is likely a better approach to experiment design than simply randomly assigning subjects to treatment and control groups.

This algorithm formalizes the two fundamental characteristics of an ideal experiment: minimal spillover and maximal similarity between control and treatment groups. If we can create algorithms to minimize both of these characteristics simultaneously and at scale, researchers can rapidly design experiments involving a massive number of subjects. This saves time while also improving the quality of the estimator of average treatment effect.

REFERENCES

Aronow, Peter M., Dean Eckles, Cyrus Samii, and Stephanie Zonszein (2020). "Spillover Effects in Experimental Data."

Sinclair, B., McConnell, M. and Green, D.P. (2012). "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments," American Journal of Political Science, 56

Baird, Sarah & Bohren, Aislinn & McIntosh, Craig & Ozler, Berk (2014). "Designing experiments to measure spillover effects," Policy Research Working Paper Series 6824, The World Bank.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. "Birds of a feather: Homophily in social networks." Annual review of sociology 27.1 (2001): 415-444.

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard?. Perspectives in biology and medicine, 48(4), 516-534.