Spillover Effects in Online Field Experiments: An Annotated Reading List

YUAN YUAN
Purdue University
and
TRACY XIAO LIU
Tsinghua University

Here we provide an overview of an important issue in online field experiments: spillover effects. We include a reading list for researchers in both academia and industry who are interested in this topic.

General Terms: Electricity, Commerce, Economics, Agents, Meta, Stuff

Additional Key Words and Phrases: Templates, Skeletons, Things

Field experiments typically aim to quantify how a given intervention (e.g., a new policy) affects certain outcomes in a population. With the growing popularity of online communities and marketplaces, there has been a corresponding increase in online field experiments. However, many online and offline experiments are subject to interference spillover effects. Spillover effects can take various forms: in some experiments, a user's outcome may be affected by the treatment assignments of other subjects; in within-subject experiments, the treatment assignment that the subject receives at one stage may affect the outcome at a later stage (also referred to as *carryover effects*).

With the presence of spillover effects, the conventional way of randomizing samples may be problematic. Intuitively, spillover effects mean that the outcome of an observation is affected not only by their own treatment assignment but also by the treatment assignments of other observations. The existence of spillover effects violates the stable unit treatment value assumption (SUTVA), the standard assumption in causal inference. For example, the existence of social contagion provides empirical evidence of the spillover effect: if one person is assigned a treatment, their family, friends, or acquaintances may also indirectly receive this treatment.

The following papers discuss how to detect or take into account spillover effects when we design experiments or analyze experimental data:

(1) Chen, Yan, and Joseph Konstan. "Online field experiments: A selective survey of methods." *Journal of the Economic Science Association* 1.1 (2015): 29-42.

    This paper presents an overview of the design and analysis of online field experiments. It covers representative studies from both economics and computer science.

Authors' addresses: yuanyuan@purdue.edu, liuxiao@sem.tsinghua.edu.cn
See also the authors' Tutorial on Experiment Design at WINE'21.

(2) Muchnik, Lev, Sinan Aral, and Sean J. Taylor. "Social influence bias: A randomized experiment." *Science* 341.6146 (2013): 647-651.

Using a randomized experiment, this study presents evidence of social influence in social networks, suggesting that the existence of spillover effects may challenge the SUTVA when we analyze experiments on social networks.

(3) Ugander, Johan, Brian Karrer, Lars Backstrom, and Jon Kleinberg. "Graph cluster randomization: Network exposure to multiple universes." In *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, pp. 329-337. 2013.

This paper designs an approach to tackle with spillover effects on large-scale social networks by creating clusters on social networks and performing randomization on the cluster level. This approach helps reduce bias when we estimate the treatment effect with the presence of spillover effects.

(4) Pouget-Abadie, Jean, Guillaume Saint-Jacques, Martin Saveski, Weitao Duan, Souvik Ghosh, Ya Xu, and Edoardo M. Airoldi. "Testing for arbitrary interference on experimentation platforms." *Biometrika* 106, No. 4 (2019): 929-940.

This paper proposes a novel approach for detecting spillover effects in social networks by simultaneously running individual (Bernoulli) and cluster level randomized experiments and comparing the resulting estimates.

(5) Holtz, David, Ruben Lobel, Inessa Liskovich, and Sinan Aral. "Reducing interference bias in online marketplace pricing experiments." *Available at SSRN* 3583836 (2020).

This paper aims to address the interference on two-sided markets. The authors also compare cluster randomization versus Bernoulli randomization and find that the latter greatly reduces bias in estimating treatment effects.

(6) Bojinov, Iavor, David Simchi-Levi, and Jinglong Zhao. "Design and analysis of switchback experiments." *Available at SSRN* 3684168 (2020).

This paper discusses the optimal design and analysis of switchback experiments. The authors also discuss how to account for carryover effects: the previous treatment assignment to one unit may affect the unit's future outcome.

(7) Aronow, Peter M. and Cyrus Samii. "Estimating average causal effects under general interference, with application to a social network experiment." *The Annals of Applied Statistics* (2017).

By relaxing SUTVA, this paper proposes a general framework, named "exposure mapping", to analyze experimental data with the presence of spillover effects.

(8) Chin, Alex. "Regression adjustments for estimating the global treatment effect in experiments with interference." *Journal of Causal Inference* 7(2), (2019).

This paper proposes regression adjustment estimators to reduce bias in experiment settings that violate SUTVA. The paper also proposes to consider the modeling of spillover effects as a feature engineering problem, which can further increase the precision of estimating treatment effects.

(9) Xu, Ya, Nanyu Chen, Addrian Fernandez, Omar Sinno, and Anmol Bhasin. "From infrastructure to culture: A/B testing challenges in large scale social networks." In *Proceedings of the International Conference on Knowledge Discovery and Data Mining* (2015). 2227–2236.

This paper discusses practices, challenges, and pitfalls in evaluating results of online controlled experiments from an industry perspective (LinkedIn).

Please note that this list is far from exhaustive and there are other related and exciting papers that are not included.