Reinforcement Learning for Respondent-Driven Sampling

Keywords: network sampling, reinforcement learning, sample design, social networks, public health

Extended Abstract

Social network data are at the forefront of healthcare research. It is widely acknowledged that a better understanding of community social structure can provide insights into the prevalence of mental illnesses such as depression, transmittable diseases such as HIV, syphilis, and COVID, and other medical conditions like obesity and type-II diabetes [1]. Currently, available methodologies for studying the underlying social network in a community fall broadly into two categories: (1) complete community census, or (2) network sampling techniques. The magnitude of modern healthcare needs makes complete censuses nearly unachievable (and generally impractical). Consequently, network sampling techniques have proven invaluable.

Respondent-driven sampling (RDS) is a network sampling algorithm based on participant referral which is frequently employed for surveillance in public health research [2]. It is especially necessary when the population of interest is hidden or hard-to-reach, e.g., people who are unhoused, people who are undocumented, and people who inject drugs. RDS begins with an initial group of individuals who are given a limited number of coupons and asked to recruit other members of the population of interest by giving them one of the coupons directly. Recipients of these initial coupons redeem them with study researchers and are then compensated, interviewed, and given new coupons to recruit additional subjects. This process continues until a sufficient sample is generated, the study budget or duration is reached, or some other stopping criterion is met.

In standard RDS, researchers give each participant an identical set of coupons. The attributes of the coupons in this set (e.g., their value, call to action, expiration date) can play a critical role in determining the coupon return rates, the coverage of the population in question, and the cost of the study. Consequently, using an identical coupon allocation throughout the study may be inefficient and ineffective in comparison to adapting the coupon allocation as more is learned about the population.

We use reinforcement learning [RL, 3] to tailor coupon allocations over time in such a way that a study objective is optimized. We consider the setting in which the goal is to recruit the largest subset of people in a hidden population with binary trait Y, e.g., undiagnosed HIV. Specifically, define the history of the RDS sampling process after recruit $n \in \mathbb{N}$ has arrived as $H^n \in \mathscr{H}^n$, a potential coupon allocation strategy as $\pi \in \Pi$, and an indicator of trait Y as $Y^v \in \{0,1\}$. We consider the value function after recruit n has arrived,

$$V^n(h^n,\pi;\beta) = \mathbb{E}_\beta \left\{ \sum_{v\geq n} Y^v(\pi) \big| H^n = h^n \right\},$$

which represents the expected future number of recruits with trait Y under policy π and the generative process, $\beta \in \mathscr{B}$. The optimal policy maximizes the value function at every $n \in \mathbb{N}$ under the true generative process, β^*, i.e., it maximizes the expected cumulative reward over the complete study.

At the beginning of the RDS sampling process, we cannot identify the optimal policy because we lack sufficient information to accurately estimate β^*. RL-RDS learns from new information to estimate the generative process and better allocate coupons. Specifically, our method balances between coupon allocations that appear to be optimal given the current estimated generative process and those that might improve estimates of β^* and thereby lead to better decisions in the future (the exploration-exploitation trade-off). This is the first principled framework for efficiently conducting RDS by directly and flexibly incorporating research objectives. We simulate RL-RDS on realistic social networks under various settings to evaluate its improvements over common network sampling techniques. Our experiments indicate that RL-RDS consistently achieves higher cumulative reward than alternatives (e.g., more efficiently recruits the target subpopulation). Figure 1 illustrates a subset of these results.

Social network data is imperative to study but difficult to collect. RL-RDS provides a natural structure for leveraging information from the network sampling process to maximize study goals. Through this work, we develop a complete, practical procedure for improving the way researchers collect representative network data.

References

[1] Xiaoyan Ma et al. "Trends in prevalence of HIV, syphilis, hepatitis C, hepatitis B, and sexual risk behavior among men who have sex with men: results of 3 consecutive respondent-driven sampling surveys in Beijing, 2004 through 2006". In: *JAIDS Journal of Acquired Immune Deficiency Syndromes* 45.5 (2007), pp. 581–587.

[2] Douglas D Heckathorn and Christopher J Cameron. "Network sampling: From snowball and multiplicity to respondent-driven sampling". In: *Annual review of sociology* 43 (2017), pp. 101–119.

[3] R.S. Sutton and A.G. Barto. *Reinforcment Learning: An Introduction*. The MIT Press, 2018.

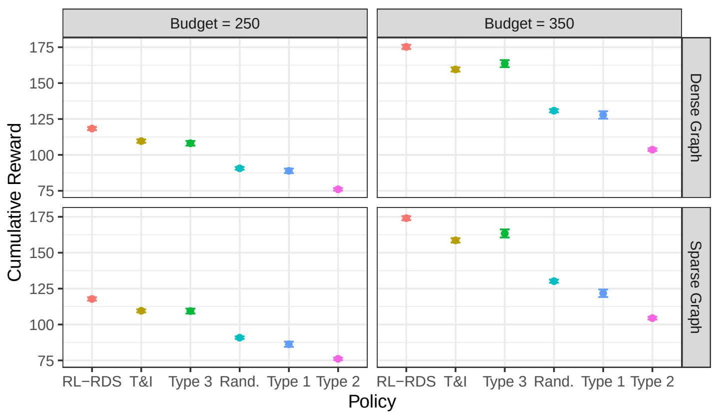

Figure 1: **Experiments.** This figure compares the estimated cumulative reward of each policy with 90% Monte Carlo confidence intervals over multiple sample budgets and graph densities. It compares RL-RDS to a random coupon allocation (Rand.), fixed coupon allocations (Type 1-3), and a two-stage design (T&I: trains a model in the first stage and implements that model to assign coupon allocations in the second stage).