Inferring Network Effects from Observational Data

David Arbour* darbour@cs.umass.edu
Dan Garant* dgarant@cs.umass.edu
David Jensen jensen@cs.umass.edu

College of Information and Computer Sciences
University of Massachusetts, Amherst

ABSTRACT

We present Relational Covariate Adjustment (RCA), a general method for estimating causal effects in relational data. Relational Covariate Adjustment is implemented through two high-level operations: identification of an adjustment set and relational regression adjustment. The former is achieved through an extension of Pearl's back-door criterion to relational domains. We demonstrate how this extended definition can be used to estimate causal effects in the presence of network interference and confounding. RCA is agnostic to functional form, and it can easily model both discrete and continuous treatments as well as estimate the effects of a wider array of network interventions than existing experimental approaches. We show that RCA can yield robust estimates of causal effects using common regression models without extensive parameter tuning. Through a series of simulation experiments on a variety of synthetic and real-world network structures, we show that causal effects estimated on observational data with RCA are nearly as accurate as those estimated from well-designed network experiments.

1. INTRODUCTION

Causal inference—estimating the effect of interventions—is central to data-driven decision making. Increased attention has been paid to randomized experimentation (A/B testing) as a method for causal inference [11, 3, 5, 6]. Recent work has extended randomized experimentation to the case of network interventions [27]. However, there are numerous circumstances where performing an experiment may be infeasible, expensive, or time-consuming.

Fortunately, a variety of methods have been devised for inferring causal effects from *observational* data. Classical methods for causal inference from observational data consist of two steps. First, an *adjustment set* [22] is identified, which consists of variables that are causally related to both the prospective cause variable (termed a *treatment*) and the potential effect variable (termed an *outcome*). Second, a procedure such as regression [21] or matching [24] is used to estimate the direct effect of treatment on outcome, correcting for the effects of the adjustment set. Extending this classical framework of estimation to relational data requires: (1) identifying adjustment sets in relational data, and (2) adjusting for the full range of the effects of those variables. Item 1 is primarily a structural question, and item 2 concerns estimation.

As a motivating example, consider the problem of estimating how a user-selected privacy setting influences the time that users spend interacting with an online social network. The privacy setting either requires users to explicitly approve others' posts to their page or it allows posting without such an approval process. Site administrators may be interested in changing the default privacy setting but want to ensure that such a change would not adversely affect site usage. Randomized experimentation on privacy settings may be controversial. Further, the propensity of users to share their posts with their friends could be influenced by characteristics of those friends. Figure 1 illustrates this example by indicating an implied correlation between social disposition and use of the privacy setting as well as a correlation between social disposition and time spent on site.

The task of adjusting for this confounding is particularly challenging because some confounding variables can be properties of neighbors in the friendship network. In Figure 1, the social disposition and privacy settings of Lucy, Sue, John, and Fred could affect both the privacy settings of Carl and the amount of time he spends on the site. The task of deciding how to set privacy policy is an intrinsically causal question because it requires reasoning about the effect that intervening on the privacy setting would have on site usage. Additionally, modeling *network effects* is of central importance—time on site is a function of the privacy settings of an entire sub-network of friends rather than the privacy setting of an individual.

In this paper, we present Relational Covariate Adjustment (RCA), the first reliable method for inferring arbitrary causal effects in networks from observational data. RCA uses a two-stage procedure. The first stage automatically identifies the set of variables that must be adjusted for. This stage uses relational d-separation [17], an extension of d-separation [22] to relational data. The second stage performs regression adjustment using relational non-parametric estimators. This adjustment procedure makes limited assumptions about the nature of the causal relationship between treatment and outcome. We provide theoretical guar-

*Equal contribution.

KDD '16, August 13-17, 2016, San Francisco, CA, USA

DOI: http://dx.doi.org/10.1145/2939672.2939791

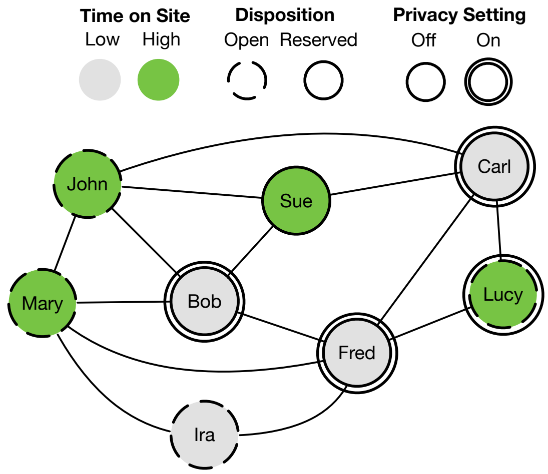


Figure 1: Social Network Privacy Example

antees showing that RCA produces a consistent estimate of causal effect.

The rest of the paper is structured as follows. Section 2 provides background for causal effect estimation and relational d-separation. Sections 3 and 4 introduce Relational Covariate Adjustment and discuss practical issues of implementation. Section 5 compares the estimates of RCA to estimates obtained via experimentation using multiple graph structures with data simulated under multiple functional forms, and shows that the performance of RCA can be competitive with experimental results.

2. PROBLEM SETUP

We assume that we are given an undirected graph $G = \langle V, E \rangle$. Let $N = |V|$, the number of vertices in the graph. Let T be a random variable composed of the treatment variables t_i of each node i in the network, so that $T = \langle t_1, t_2, \ldots, t_N \rangle$. Let $\boldsymbol{\pi}$ be an assignment to T, that is, $\boldsymbol{\pi} = \langle \pi_1, \pi_2, \ldots, \pi_N \rangle$, where π_i is an assignment to t_i. The *average causal effect* (ACE) is defined as the expected difference in outcome Y under treatment $\boldsymbol{\pi}$, contrasted with an alternate treatment $\boldsymbol{\pi}'$:

$$\text{ACE}(\boldsymbol{\pi}, \boldsymbol{\pi}') = E[Y|do(T = \boldsymbol{\pi})] - E[Y|do(T = \boldsymbol{\pi}')]. \quad (1)$$

Throughout the paper, we use the *do* operator [22] to refer to the interventional distribution, that is, the distribution that would arise due to manipulation of T rather than passive observation. Equation 1 may also be expressed in the potential outcomes framework [25] by regarding Y as a node-specific function of treatment. Ugander et al. [27] consider a special case of equation 1 where $\boldsymbol{\pi} = \vec{1}$ and $\boldsymbol{\pi}' = \vec{0}$. Hudgens and Halloran [9] refer to the above quantity as the *population average overall causal effect.*

Expressed directly in equation 1 is the notion that the outcome of subject i is a function of the entire treatment assignment vector, not only π_i. This distinction is critical for estimating network effects, as we now have a language to express interventions on multiple subjects. When dealing with causal quantities as in equation 1, it is common to assume that $E[Y|do(T = \boldsymbol{\pi})]$ is invariant with respect to treatment assignments to nodes which do not neighbor

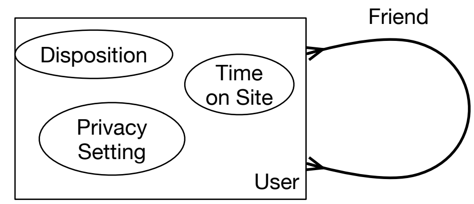


Figure 2: ER Diagram for Social Network Example

i. Let T_{nbr_i} denote the treatment variables of i's neighbors, and let $\boldsymbol{\pi}_{nbr_i} = \{\pi_j | \{i,j\} \in E\}$ and $\boldsymbol{\pi}'_{nbr_i}$ be multisets representing assignments to T_{nbr_i}. The neighborhood invariance assumption leads to the following reformulation of the average causal effect:

$$\text{ACE}(\boldsymbol{\pi}, \boldsymbol{\pi}') = \frac{1}{N} \sum_{i=1}^{N} E[Y|do(t_i = \pi_i, T_{nbr_i} = \boldsymbol{\pi}'_{nbr_i})] \\ - E[Y|do(t_i = \pi'_i, T_{nbr_i} = \boldsymbol{\pi}'_{nbr_i})]. \quad (2)$$

Equation 2 is consistent with the peer exposure models considered by Aronow et al. [2], Toulis and Kao [26], and the notion of effective treatments considered by Manski [18]. These causal quantities facilitate answering questions about interventional strategies including:

1. $E[Y|do(t_i = 1, T_{nbr_i} = \vec{1})] - E[Y|do(t_i = 0, T_{nbr_i} = \vec{0})]$: How would individual i's outcome change if i and its neighborhood were to be treated, as opposed to untreated? This quantity is the basis of ACE$(\vec{1}, \vec{0})$, the quantity considered by Ugander et al. [27] and Gui et al. [8].

2. $E[Y|do(t_i = 1, T_{nbr_i} = \vec{0})] - E[Y|do(t_i = 0, T_{nbr_i} = \vec{0})]$: How does subject i's expected outcome change if i is treated but no neighbors are treated? We might think of this effect as an "insulated" individual effect.

3. $E[Y|do(t_i = 0, T_{nbr_i} = \vec{1})] - E[Y|do(t_i = 0, T_{nbr_i} = \vec{0})]$: How does subject i's expected outcome change if i is left untreated but all neighbors are treated?

By considering different settings of $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$, we can examine a large number of possible intervention strategies, without being restricted to applying the same "type" of intervention to each node in the network. In practice, no single value of $\boldsymbol{\pi}$ could be used to apply interventions (2) and (3) in the list above to all nodes in the network. However, we can consider targeted interventions on specific individuals in the network, so it is useful to consider these effects.

3. RELATIONAL ADJUSTMENT SETS

We now briefly introduce the relational concepts necessary to describe Relational Covariate Adjustment, following the notation and terminology of Maier et al. [17, 16].

3.1 Relational Causal Graphical Models

Let a *relational schema* $\mathcal{S} = (\mathcal{E}, \mathcal{R}, \mathcal{A}, card)$ be the set of entity, relationship, and attribute classes of a domain. It includes a cardinality function that imposes constraints on the number of times an entity instance can participate in a relationship. Without loss of generality, we will focus our presentation on the case of a simple network, where there is a single entity, and a single many-to-many relationship, e.g. a social network. Continuing the example of Figure 1:

$$\mathcal{E} = \{\text{Users}\},$$
$$\mathcal{R} = \{\text{Friend}\},$$
$$\mathcal{A} = \{\text{time on site, disposition, privacy setting}\},$$
$$card(\text{Connected}) = \text{Many}.$$

Users are connected to potentially many other users, each of which has a time on site, disposition, and privacy setting attribute. Relational schemas are often visualized with entity-relationship diagrams as in Figure 2.

A *relational skeleton* is a partial instantiation of a relational schema that specifies the set of entity and relationship instances that exist in the domain. Using our online social network example, this corresponds to specific users and the friends that they connect to through the site. With a given schema, a *relational path* can be defined, which is a predicate that defines a path with respect to a schema. In our example, relational paths correspond to friendship paths, defined through the connectivity properties of the online social network. We will refer to variables with a trivial relational path (e.g., the immediate attributes of individuals), as *propositional variables. Relational variables* consist of a relational path and an attribute that can be reached through that path. For instance, the multiset of privacy settings for friends adjacent to user i is a relational variable. Relational variables can have causal dependencies defined between them, specified by a *relational model* $\mathcal{M} = (\mathcal{S}, \mathcal{D})$. This model consists of a collection of relational dependencies ($\mathcal{D}$) defined over a relational schema ($\mathcal{S}$). The relational model represents, as one example, the property that a user's time on site is affected by the privacy settings of adjacent users. $\mathcal{M}$ also specifies a parametrized conditional distribution of each relational variable given its parents. In the context of this work, we do not have access to these distributions and must estimate them from data.

To evaluate conditional independence queries on a model $\mathcal{M}$, we first construct an abstract ground graph (AGG) [17], a lifted representation that admits the computation of d-separation queries on multi-relational domains. Abstract ground graphs are defined from a given perspective, specifying a *base item* of the analysis, and include nodes that correspond to relational variables. In general, the construction of an AGG can involve creating auxiliary "intersection" variables. However, for the case of single-entity, single-relationship networks (e.g. social networks or simple communication networks) there exists a single AGG that can be represented without the use of auxiliary variables. That is:

Proposition 1. *Given a model with a single entity single, relationship schema, the complete set of d-separation facts can be determined by considering only propositional variables and relational variables.*

The proof stems from a direct application of relational d-separation and is presented in the Appendix. The consequence of Proposition 1 is a relatively simple representation[1]. Within our running example there is a single perspective (person) and relational variables are defined with respect to the relative distance to an individual (e.g. friends and friends of friends). One plausible abstract ground graph for this example is shown in Figure 3, in which the disposition and privacy settings of a person and her friends affect her time spent on site. Note that in Figure 3 there are two different types of variables present. Propositional variables are those preceded by U^0, and are measured on a single instance. Relational variables are named as U^i, for $i > 0$. These variables representing the values of a person's friends and the friends of her friends, respectively. Given the AGG, conditional independence facts can be computed directly using the same rules of d-separation used for Bayesian networks. For instance, from Figure 3, we can see that $U^2.Prv \perp\!\!\!\perp U^0.Prv | U^1.Prv$, because $U^1.Prv$ blocks all d-connecting pathways between the privacy settings of *friends of friends* and a user's time spent on site. These d-separation properties are essential to identifying a sufficient set of conditioning variables for a given causal query, discussed in more detail in the following section.

[1]Relative to those required for multi-relational domains.

Notation	Meaning
$U_i^0.ToS$	The value of variable ToS for instance i of entity U. For instance, this could represent the time on site of user i.
$U_i^1.\mathbf{ToS}$	A multiset representing the value of variable ToS on instances related to instance i of entity U through a path of length 1. For instance, this could represent the time that friends of user i spend on the site. We can represent users that are friends with i's friends with the notation U_i^2, and so on.

Table 1: Relational Notation

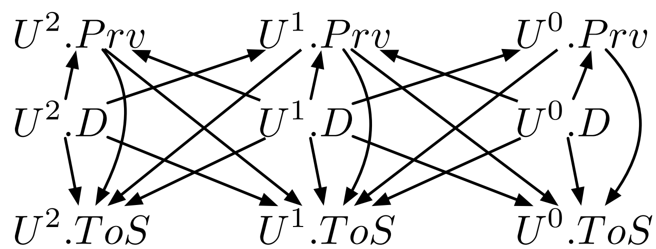


Figure 3: Abstract Ground Graph for the Social Network Example. In this example, each user's disposition ($U^0.D$) affects that user's privacy settings ($U^0.Prv$) and time on site ($U^0.ToS$). Further, the dispositions and privacy settings of a user's immediate peers ($U^1.D$ and $U^1.D$, respectively) affect that user's time on site. A user's privacy settings are also influenced by their peers' privacy settings. This structure repeats for U^2, representing friends of friends. Higher orders of U^p can be considered, but are not shown here.

3.2 Relational Backdoor Criterion

With a suitable representation in hand, we now turn to the core aim of this work: identifying interventional distributions. The approach taken here is to use an extension of the *back-door criterion* [22] to relational domains:

Definition 1. *(Relational Back-Door Criterion) A set of variables* $\mathbf{C}$ *satisfies the* relational back-door criterion *with* *respect to variable sets* (X_1, X_2) *in an AGG G if:*

1. *No node in* $\mathbf{C}$ *is a descendant of any node in* X_1 *in the AGG (equivalently, no node in* $\mathbf{C}$ *is a post-treatment variable); and*

2. $\mathbf{C}$ *blocks every back-door path between* X_1 *and* X_2 *in the AGG*

Note that here a *back-door path* refers to a path with an arrow into a member of X_1. Definition 1 is a direct extension to relational data of the back-door criterion presented by Pearl [22]. In the case of a single entity with no relationships, the definition reduces to the propositional case.

When such a set $\mathbf{C}$ can be identified, an estimate of the interventional distribution can be obtained through a simple application of the adjustment formula:

$$P(X_2|do(X_1 = x)) = \int_c P(Y|X_1 = x, \mathbf{C} = c)\mathrm{d}P(\mathbf{C} = c) \quad (3)$$

Then, average causal effects can be computed as follows:

$$ACE = E[X_2|do(X_1 = x)] - E[X_2|do(X_1 = x')] \quad (4)$$
$$= \int_c yP(X_2|X_1 = x, \mathbf{C} = c)\mathrm{d}P(\mathbf{C} = c)$$
$$- \int_c yP(X_2|X_1 = x', \mathbf{C} = c)\mathrm{d}P(\mathbf{C} = c), \quad (5)$$

where P represent either a probability density or probability mass function. Semantically, because relational variables take on values that may be multisets, there is a notion of *exhangeability* encoded in this estimation framework. Consider once again the example of Figure 1. In this case, Sue has three neighbors, John, Bob and Carl. Let $U^1_{\text{Sue}}.\mathbf{Prv}$ represent the multiset of time on site values of these neighbors (see Table 1). As presented, $U^1_{\text{Sue}}.\mathbf{Prv}$ takes on the value {On, On, Off}. Intervention on Carl or Bob's privacy setting would yield the interventional regime $do(U^1_{\text{Sue}}.\mathbf{Prv} =$ {On, Off, Off}). As such, our interventional language is invariant with respect to the *identities* of the instances under intervention, and focuses strictly on the variables measurable on those entities.

3.2.1 Connection to Network Experimentation

There is a close relationship between Relational Covariate Adjustment and the adjustments performed for peer-effects in the network experimentation literature (c.f., [2, 27, 8]). Given this connection, we discuss this relationship for readers familiar with network experimentation. Current work in network experimentation are described within the potential-outcomes framework and assume *strong ignorability*, i.e., that (1) the outcome is rendered independent of treatment given treatment status and (2) that all instances have a treatment probability, $p \in (0, 1)$. Within non-network experiments condition (1) is trivially satisfied via randomization. However, even in the simple network setting there is dependence between other treatments and an individual's outcome. Further, by virtue of network randomization designs (i.e., [27, 8]), dependence is induced between the treatment status of instances. This dependence is depicted in Figure 4. The graphical view shows that simple use of Relational Covariate Adjustment can be applied to adjust for network bias, with $U^1.Prv$ constituting the adjustment set. Thus, the estimator of Gui et al. [8] can be seen as a special case of Relational Covariate Adjustment, with an assumed dependence structure of Figure 4 and adjustment performed with a linear model. However, in contrast to current network experimentation estimation methods, Relational Covariate Adjustment can be applied easily to observational data with multi-valued and continuous treatments and an arbitrary number of confounders without modification.

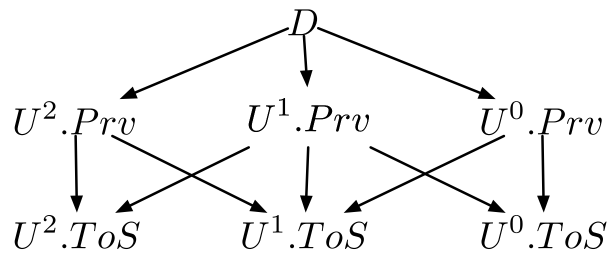


Figure 4: An abstract ground graph representing the dependence structure under network experiment. This structure is similar to 3, except that disposition no longer influences privacy settings, and is excluded from the diagram. A variable D representing the experimental design may induce marginal dependence between treatments. It is possible that the outcome of peers ($U^1.ToS$) affects $U^0.ToS$, but including $U^1.Prv$ in a conditioning set is sufficient to satisfy the back-door criterion for treatment $U^0.Prv$.

4. EMPIRICAL ESTIMATION

We now discuss how to practically estimate the effects of interventions in relational domains. In contrast to the non-relational setting, computing the adjustment formula in equation 5 is not straightforward because the hypothetical values of X_1 could be multisets. We present a strategy for conditioning on multisets that does not make strong assumptions about functional form. Algorithm 1 presents the procedure. Step 1 identifies the adjustment set by using relational d-seperation to find the necessary set of variables $\mathbf{C}$ to block all back-door paths between T and Y.

The causal effect is then estimated as

$$\mathbb{E}[Y|do(T = t)] = \int_C yP(Y = y|T = t, \mathbf{C} = c)\mathrm{d}P(\mathbf{C} = c) \quad (6)$$

$$\approx \frac{1}{N}\sum_{i=1}^{N} y_i P(Y = y_i|T = t, \mathbf{C} = c_i) \quad (7)$$

$$= \frac{1}{N}\sum_{i=1}^{N} E[Y|T = t, \mathbf{C} = c_i], \quad (8)$$

where equation 7 is a Monte-Carlo approximation to the integral. $E[Y|T = t, \mathbf{C} = c_i]$ can be estimated from a regression of y on features T and $\mathbf{C}$.

4.1 Calculating Network Effects

Algorithm 1 can be applied to estimate a variety of causal effects derived from the definition presented in equation 2. In what follows, $U^0.T$ refers to a subject's treatment and $U^1.T$ refers to the treatments of immediate neighbors.

Algorithm 1: `RelationalAdjustment`

Input: Relational model $\mathcal{M}$, outcome Y, treatment(s) X

Output: $h(x) = \sum_{i=1}^{N} E[Y|do(X = x)]$

Use relational d-separation to identify adjustment set $\mathbf{C}$ for causal effect of X on Y

Estimate $E[Y|X, \mathbf{C}]$ via regression or classification

$h(x) = \sum_{i=1}^{N} E[Y|X = x, \mathbf{C} = \mathbf{c}_i]$

return $h(x)$

Marginal Individual Effect

$$\begin{aligned} h &= \texttt{RelationalAdjustment}(\mathcal{M}, U^0.Y, U^0.T) \\ h(1)-h(0) &= E[U^0.Y|do(U^0.T=1)] \\ &\quad - E[U^0.Y|do(U^0.T=0)] \end{aligned} \tag{9}$$

This effect represents the expected change in an arbitrary subject's outcome, $U^0.Y$, when considering two alternate settings of that subject's treatment, $U^0.T$ (1 and 0). The function h represents the expected outcome when applying a hypothetical intervention to $U^0.T$, conditioning on $\mathbf{C}$. Additionally, the treatment assignment of peers, $U^1.T$, can influence both $U^0.T$ and $U^0.Y$, which requires including peer treatment values in the set of confounders, i.e, $U^1.T \in \mathbf{C}$.

Marginal Peer Effect

$$\begin{aligned} h &= \texttt{RelationalAdjustment}(\mathcal{M}, Y, U^1.T) \\ h(\theta)-h(\theta') &= E[U^0.Y|do(U^1.T=\theta)] \\ &\quad - E[U^0.Y|do(U^0.T=\theta')] \end{aligned} \tag{10}$$

The above case concerns the causal effect of settings of the treatment assignments of peers, $U^1.T$. θ and θ' are multisets consisting of the treatment values of neighbors. For instance, in the context of Figure 1, $\theta_{\mathrm{Sue}} = \{\mathrm{On}, \mathrm{On}, \mathrm{Off}\}$. We could consider altering the treatment of Sue's neighborhood to $\theta'_{\mathrm{Sue}} = \{\mathrm{Off}, \mathrm{Off}, \mathrm{On}\}$. The effect of the intervention is given by $h(\theta_{\mathrm{Sue}}) - h(\theta'_{\mathrm{Sue}})$. This formulation facilitates the estimation of arbitrary treatment settings of a node's neighborhood.

Total Effect

$$\begin{aligned} h &= \texttt{RelationalAdjustment}\big(\mathcal{M}, Y, (U^0.T, U^1.T)\big) \\ h(1, \vec{1}) - h(0, \vec{0}) &= E[U^0.Y|do(U^0.T=1, U^1.T=\vec{1})] \\ &\quad - E[U^0.Y|do(U^0.T=0, U^1.T=\vec{0})] \end{aligned} \tag{11}$$

This effect represents an intervention on both $U^0.T$ and $U^1.T$. The adjustment procedure is valid for simultaneous interventions on these variables because the back-door criterion (Definition 1) applies to *sets* of variables. Now, h is a function of two variables, the hypothetical intervention to $U^0.T$ and the hypothetical intervention to $U^1.T$. The first argument to h is, in the case of binary treatments, 0 or 1. The second argument to h is a multiset. This class of effects is most applicable to estimation of applying an intervention to all individuals on a network, e.g., a site-wide feature rollout.

4.2 Summarizing Relational Features

When any term in the adjustment equation is a relational variable, $E[Y|T, \mathbf{C}]$ cannot be directly estimated using regression or classification estimators designed for independent and identically distributed data because relational variables' instances consist of multisets rather than single observations. A common approach to address this is to create aggregations to succinctly represent the sets with a small number of real-valued features. There is a long history in statistical relational learning of using user-specified aggregation functions to model the distribution of a relational variable [12, 23]. While these approaches have yielded impressive results for the task of prediction, causal inference requires stronger guarantees about what is being captured by the aggregation functions. The aggregation function should be a sufficient statistic of the underlying distribution of the variable, rendering model parameters independent of the data. For instance, specifying the `mean` aggregation would be sufficient if the values of a relational variable are Poisson distributed, and in the case of a normal distribution, the `variance` aggregation must also be present. When sufficient statistics are employed, then we can be confident that *all relevant aspects* of the distribution of a set have been accounted for when marginalizing to compute the interventional distribution. When assumptions can be made about the marginal distribution of relational variables, a set of features can be constructed for regression by taking the sufficient statistics for each instance of a relational variable. Once this set is constructed, any consistent regression or classification model can be used to estimate $E[Y|T, \mathbf{C}]$. In the absence of known sufficient statistics, estimates of a number of the moments of a distribution can be used as an approximate solution. We assume that the sufficient statistics S_x of the true distribution can be described as a function of its k-th order moments:

$$S_x = f(M_1(X), \ldots, M_k(X))$$

where $M_k(X) = \frac{\sum_i^N X_i^k}{N} \approx \int x^k \hat{p}(x)dx$ is the empirical estimate of the k-th moment of X. This implies the following procedure: (1) for each relational variable generate a set of k aggregates of the $1, \ldots, k$ moments of the set, (2) use this new data set as the features to a non-linear regression or classification model to estimate $E[Y|T, \mathbf{C}]$.

5. EXPERIMENTS

In this section we evaluate whether, and under which circumstances, Relational Covariate Adjustment can serve as a feasible alternative to experimentation for causal inference. To that end, we constructed an evaluation suite to compare RCA to state-of-the-art techniques for estimating causal effects from experiments. We provided experimental techniques with *experimental* data, and we provided RCA-data with more challenging data sets in which relational confounding variables are present. We examined a variety of real and synthetic networks, using simulated data with multiple functional relationships between treatment and outcome.

5.1 Synthetic Data Generation

Data generation process was performed as follows:

1. Generate a random network

2. Sample treatment using one of two regimes:
   (a) `Exp`: Sample treatment from an experimental context, in which treatment is assigned using a graph clustering technique
   (b) `Obs`: Sample treatment as a function of confounding variables and possibly treatments of neighbors in the network
3. Sample outcome according to the treatment assigned in step (2). In the `Obs` regime, outcome is a function of confounding variables and treatment. In the `Exp` regime, outcome is a function of treatment assignments.

In both the `Obs` regime and the `Exp` regime, the task is identical: estimate the relationship between treatments (individual and those of peers) and outcomes. We compared the performance of models learned from the observational data to estimates obtained by experimentation[2].

5.1.1 Synthetic networks

We considered two network structures in our synthetic experiments: small-world networks and preferential attachment networks. For small-world networks, each node has degree (in+out) of 10 in the initial lattice. We varied the rewiring probability in $\{0, 0.01, 0.1, 0.15\}$. A rewiring probability of 0 results in a regular lattice, and a rewiring probability of 1 results in a random (Erdős-Rényi) network. For preferential attachment networks, we varied the power of the attachment in $\{0.1, 0.5, 1\}$. In all cases, the synthetic networks we consider have 1024 nodes.

Each network has a simple relational model consisting of a single entity (U) and relationship (adjacency). Each instance of U (i.e., a node in the network) has four attributes, C_1, C_2, T, and Y. We are interested in estimating the effects of $U^0.T$ (intrinsic treatment) and $U^1.T$ (treatment of peers) on $U^0.Y$ (intrinsic outcome).

5.1.2 Treatment Models

In the `Obs` regime, propensity for treatment can be caused by intrinsic covariates, covariates of peers, and treatments of peers. To simulate data from that regime, we first constructed a confounding term L_i which is a linear combination of:

- $U^0.C_1$
- $U^0.C_2$
- $\texttt{mean}(U^1.C_1)$
- $\texttt{mean}(U^1.C_2)$
- $\texttt{var}(U^1.C_1)$
- $\texttt{var}(U^1.C_2)$
- $\texttt{mean}(U^1.C_1) * \texttt{var}(U^1.C_1)$
- $\texttt{mean}(U^1.C_2) * \texttt{var}(U^1.C_2)$

Then, treatment is sampled as a binomial random variable with success probability that is a logistic function of L_i. To simulate influence between the treatments of subject i and its neighbors, we use a Gibbs sampling technique inspired by Manski [18]. After initially assigning treatment, we resample treatment with an additional parameter $\theta_{nbr_i,s-1}$, the proportion of i's neighbors that are treated at the previous iteration. This process is repeated until $s = 3$.

$$T_{i,0} \sim \textit{Binom}\left(\textit{logistic}\left(\beta_L L_i + \epsilon\right)\right) \tag{12}$$

$$T_{i,s} \sim \textit{Binom}\left(\textit{logistic}\left(\beta_L L_i + \beta_T \theta_{nbr_i,s-1} + \epsilon\right)\right) \tag{13}$$

Here, $\epsilon \sim \mathcal{N}(0, 1)$. We vary the strength of the confounding coefficient, β_L, from 0 to 3. We vary the strength of dependence on peers' treatments, β_T, from 0 to 10. When $\beta_T = 2$, we find that the distribution of peer treatment proportions, θ_{nbr_i}, is roughly uniform. When $\beta_T = 10$, this distribution is bi-modal with peaks at 0 and 1.

In the `Exp` regime, treatment was assigned randomly (with probability 0.5) at the level of *graph clusters* rather than individuals using a technique outlined by Ugander *et al.* [27]. This clustering technique assigns treatment in such a way that nodes are more likely to have completely treated or completely untreated neighborhoods. In other words, graph cluster randomization leads to bi-modal distributions of θ_{nbr_i} with peaks at 0 and 1. This randomization technique is employed by experimental estimators to estimate the total effect of equation 11.

5.1.3 Outcome Models

We explored the use of three distinct outcome forms. In the first case, outcome is a linear function of individual treatment, the proportion of treated peers, θ_{adj_i}, confounding variables L_i, with noise that is distributed as a standard normal. The general form of this function is shown below in equation 14. The relationship between the θ_{adj_i} and outcome is shown in Figure 5b for a specific parameter setting.

$$Y_i \sim \beta_I T_i + \beta_P \theta_{nbr_i} + \beta_L L_i + \epsilon \tag{14}$$

We also considered non-linear functions of treatment and covariates. The first of these is shown in equation 15, and is a sigmoid function of T_i, θ_{nbr_i}, and L_i. Figure 5a shows one instance of this function class. This function is bounded in the range $(0, 1)$. As β_I and β_P grow, the outcome approaches 1 more sharply.

$$Y_i \sim \left(1 + \exp\left(-\left(2\beta_I T_i + 2\beta_P \theta_{nbr_i} + \beta_L L_i + \epsilon\right)\right)\right)^{-1} \tag{15}$$

The final outcome model we use is linear in T_i and L_i, but depends on θ_{nbr_i} through a radial basis function about 0.5. An instance of this function can be seen in Figure 5c. In this case, the outcome peaks when $\theta_{nbr_i} = 0.5$.

$$Y_i \sim \beta_I T_i + \exp\left(-\left(\beta_P \theta_{nbr_i} - 0.5\right)^2\right) + \beta_L L_i + \epsilon \tag{16}$$

In what follows, we refer to the functions outlined in equations 14, 15, and 16 as *linear*, *sigmoid*, and *RBF*, respectively.

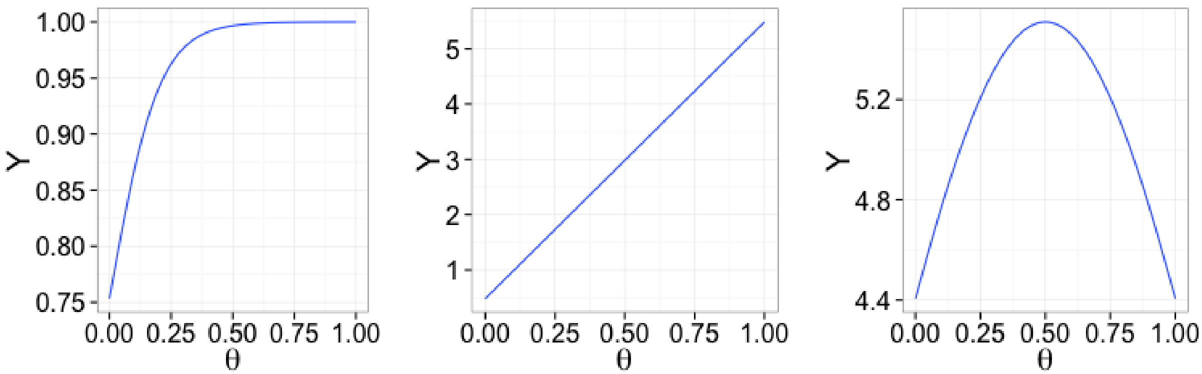
5.2 Estimators

In practice, any consistent conditional estimator of $E[Y|T, \theta_{nbr_i}, \mathbf{C}]$ will satisfy the requirements of the relational adjustment technique, provided $\mathbf{C}$ satisfies the relational back-door criterion. For our experiments, we used gradient boosted trees (GBMs) to model this expectation, where $\mathbf{C}$ consists of the means and variances of $U^1.C_1$ and $U^1.C_2$.

Gradient boosted trees [7] are a nonparametric ensemble where each base learner is a low-depth decision tree. At each iteration training samples are reweighted according to their predictive error on the previous iteration. The boosting procedure has been shown to be consistent [7], and provides near state-of-the-art results on a variety of tasks.

We employed two experimental effect estimators within the `Exp` regime, Horvitz-Thompson estimation [27] and a linear additive model [8]. The Horvitz-Thompson estimator can be written as a weighted sum of outcomes of nodes which fall into two distinct exposure categories. We defined a node i as "exposed" if $T_i = 1$ and $\theta_{nbr_i} > 0.75$. We defined a node

[2]Code used to reproduce these experiments is available at https://github.com/darbour/RelationalAdjustment



(a) Sigmoid, $\beta_P = 10$ (b) Linear, $\beta_P = 5$ (c) RBF, $\beta_P = 5$

Figure 5: Examples of outcome models considered in this work, shown here as a function of the proportion of treated friends.

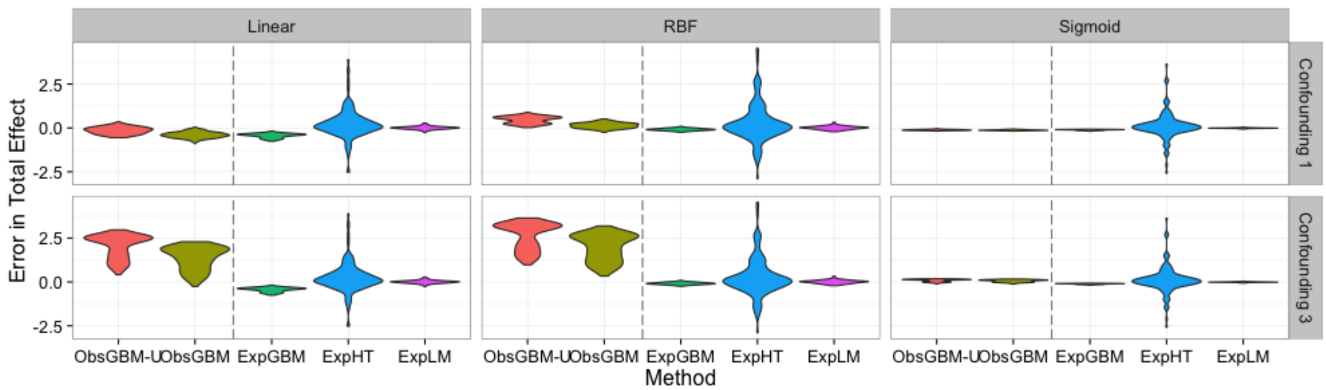


Figure 6: Accuracy of experimental and observational effect estimates across various outcome models as confounding strength is varied.

as "non-exposed" if $T_i = 0$ and $\theta_{nbr_i} < 0.25$. Nodes which do not fall into one of these categories are not used in the estimation process.

$$\frac{1}{N}\sum_{i=1}^{N}\frac{Y_i\mathbb{I}(\theta_{nbr_i} > 0.75, T_i = 1)}{P(\theta_{nbr_i} > 0.75, T_i = 1)} - \frac{Y_i\mathbb{I}(\theta_{nbr_i} < 0.25, T_i = 0)}{P(\theta_{nbr_i} < 0.25, T_i = 0)}.$$

The probabilities in the denominator are estimated using the dynamic programming algorithm introduced by Ugander et al. [27]. This method is useful primarily when the effect of interest is the total effect and the distribution of θ_{nbr_i} is bimodal with peaks at 0 and 1. We refer to this estimation strategy as `ExpHT`.

The linear additive model introduced by Gui et al. [8] fits the conditional expectation $E[Y|T, \theta_{nbr_i}]$, which is appropriate when treatment is assigned experimentally and outcome is a linear model. We refer to this model as `ExpLM`.

It is important to note that, for both `ExpHT` and `ExpLM`, the results reported are with respect to a performed experiment. This is contrast to the setting of Relational Covariate Adjustment, which is given access *only* to observational data, without the benefit of randomization.

5.3 Findings

For each combination of parameter settings, spanning rewiring probability, β_I (individual effect), β_P (peer effect), β_L (confounding strength), and β_T (treatment auto-correlation), we performed 25 trials to assess variance. This resulted in $2269 \times 25 = 56{,}725$ datasets, some belonging to the experimental regime (`Exp`) and some belonging to the observational regime (`Obs`). Within the experimental regime, we estimated total effect using `ExpHT` and `ExpLM`, the current state-of-the-art techniques for effect estimation on networks. Within the observational regime, we used the Relational Covariate Adjustment procedure with gradient boosted trees. This is referred to as `ObsGBM` in what follows. We also include results for an unadjusted GBM which does not include any relational covariates. We refer to this model as `ObsGBM-U`.

Accuracy of the total causal effect estimates for the linear outcome model are shown in Figure 7. Each box in this figure represents the distribution of performance values across all network settings and model parameterizations. These results indicate that the `ExpLM` model performs best in this context, with `ExpHT` yielding slightly more bias and significantly more variance. However, in the observational case, which is a more complex estimation task, the Relational Covariate Adjustment implementation, `ObsGBM` is competitive with the linear model in terms of bias, and yields only slightly more variance than the HT estimator.

We also examined the performance of these models across a variety of outcome functions. The error in total causal ef-

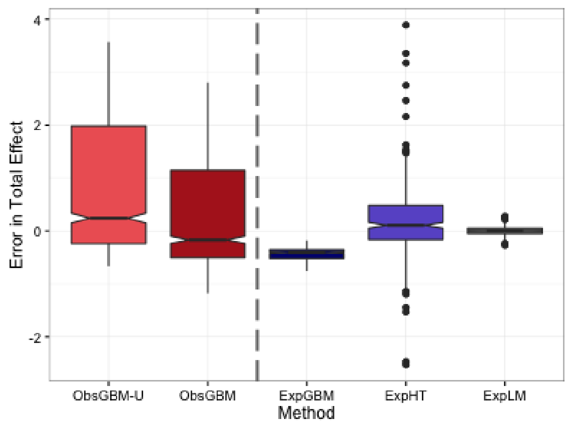


Figure 7: Comparison of estimates obtained from retrospective, confounded, observational data (left) and those from experimentation (right). An *overestimated* effect results in positive error, and an *underestimated* effect results in negative error. These methods almost always overestimate the true global effect.

	Exp. LM	Obs. GLM
Linear	**0.0869 (0.0752)**	0.6527 (0.5936)
RBF	**0.102 (0.0877)**	0.2342 (0.1687)
Sigmoid	**0.0178 (0.0158)**	0.0269 (0.0157)

Table 2: Root mean squared error for marginal individual effects. One standard error is shown in parentheses.

fect estimates are shown in Figure 6. This demonstrates two dimensions of variability in our simulations. First, different functional forms lead to more or less challenging estimation tasks. Most significantly, as the strength of confounding (β_L) is increased from 1 to 3, the observational regime becomes more challenging. This matches intuition—in the extreme, where $\beta_L = 0$, any confounding between treatment and outcome disappears.

While the `ExpHT` model is specifically designed to estimate only total effects, `ExpLM` and `ObsGBM` can also compute marginal individual effects and marginal peer effects. We computed the root mean squared error between estimated individual effects and actual individual effects—this error is shown in Table 2. `ObsGBM` is competitive with `ExpLM` primarily for non-linear functional forms.

Finally, we examined the ability of the `ExpLM` and `ObsGBM` to model marginal peer effects. Unlike the total effect and the marginal individual effect, there is a *spectrum* of peer effects induced by varying θ_{nbr_i}. The possible functional relationships between θ_{nbr_i} and Y_i are shown in Figure 5. Figure 8 provides another concrete example of such a function along with the models estimated by `ExpLM` and `ObsGBM`.

	Exp. LM	Obs. GBM
Linear	**0.0535 (0.021)**	0.6241 (0.485)
RBF	0.4476 (0.248)	**0.403 (0.264)**
Sigmoid	0.0661 (0.015)	**0.0391 (0.027)**

Table 3: Root mean squared error for marginal peer effects. One standard error is shown in parentheses.

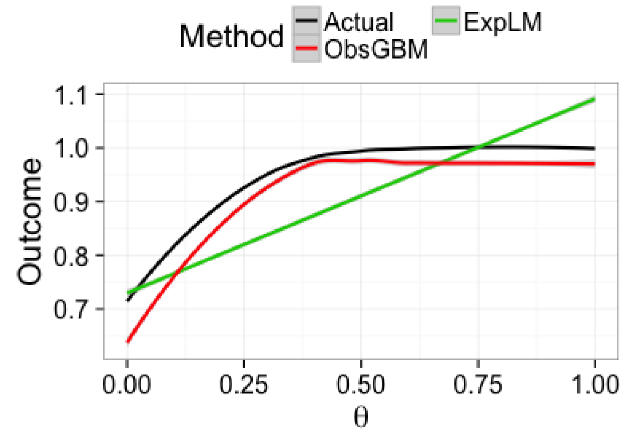


Figure 8: An example of the sigmoid outcome model. In this case, a model of the marginal peer effect is estimated from observational data using a boosted model and from experimental data with a linear model, with $\beta_I = \beta_P = 5$ and $\beta_L = 1$.

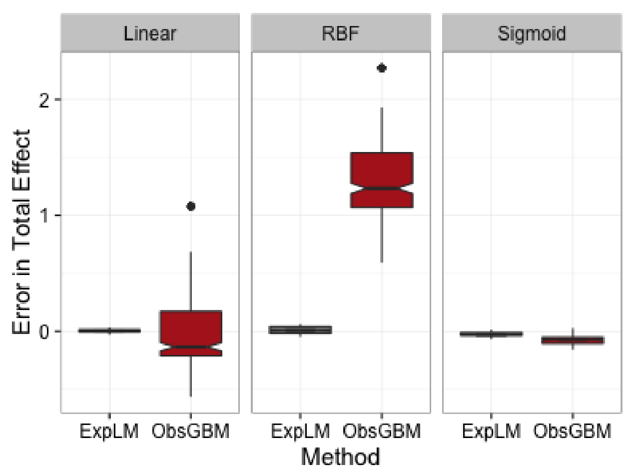


Figure 9: Estimated Total Effects in the Enron Data

Table 3 shows the root mean squared error for peer effects $\theta_{nbr_i} \in \{0, 0.1, \ldots, 0.9, 1\}$. In the linear case, the `ExpLM` model has an advantage over `ObsGBM`. However, the importance of modeling non-linearity becomes clear in the RBF and sigmoid examples, for which the observational estimator is superior to the experimental estimator.

5.4 Real Networks

To demonstrate the applicability of RCA to large networks for which the edge generation process is unknown, we also compared the performance of `ObsGBM` and `ExpLM` on the Enron graph [14]. The nodes of the Enron network are individuals, with an edge between them if either of them have sent an email to the other. The network is contains 36,692 nodes and 183,831 edges in total, with a clustering coefficient of approximately 0.5 and diameter of 11. In the absence of ground truth measures, we generated synthetic random variables following the procedure of Section 5.1, and use the real

	ExpLM	ObsGBM
Linear	**0.0399 (0.0279)**	1.4038 (0.3772)
RBF	**0.6242 (0.2571)**	1.2883 (0.3778)
Sigmoid	0.3439 (0.2113)	**0.0561 (0.0101)**

Table 4: Root mean squared error for marginal individual effects in Enron data.

	ExpLM	ObsGBM
Linear	**0.0213 (0.007)**	0.4855 (0.195)
RBF	0.5255 (0.148)	**0.2278 (0.112)**
Sigmoid	0.2703 (0.14)	**0.0266 (0.025)**

Table 5: Root mean squared error for marginal peer effects in Enron data.

graph topology to test scalability and efficacy. The form of the generative functions were identical to those used in the observational setting. We then measured the estimates of the total effect, marginal peer effects, marginal individual effects for each method.

Figure 9 shows the quality of estimated total effects across each outcome model. While the results from a synthesized network experiment are superior to the estimates from ObsGBM, the results are on comparable scales. Table 4 and Table 5 show the error in estimated individual and peer effects, respectively. Again, the results from ObsGBM are similar to ExpLM. ExpLM performs exceptionally well at experimental data with a linear outcome. However, ObsGBM has a clear advantage in estimating marginal peer effects under the RBF and Sigmoid models. We conjecture that the scale-free nature of the Enron network leads to particularly favorable circumstances for experimental approaches such as ExpLM. Scale-free networks have many nodes with only one or two neighbors, thus the probability that an entire neighborhood will be completely treated or completely untreated is relatively high.

6. RELATED WORK

Ugander et al. [27] and Gui et al. [8] present methods which aim to measure the effect of placing the entire network under treatment versus control. In both cases, this is achieved by partitioning the graph into clusters, treating each cluster randomly and estimating the the average causal effect after adjustment for peer confounding. Toulis and Kao [26] consider experimental design and estimation to measure average peer effect as a quantity of interest. In both cases, the methods discussed within this paper can be seen as complimentary work, providing interpretation within the causal graphical models framework. This interpretation aids the identification of threats to validity and provides a unified framework for estimation of a variety of causal effects. Importantly, the causal graphical model view of this work also admits inference in the non-experimental setting.

There have been numerous applications in recent year that seek to measure causal effects in real relational domains. Bakshy et al. [4] performed large scale experiments to understand the effect of social cues on consumers' receptiveness to advertisements. Aral and Walker [1] used experimentation to understand the process of social diffusion, or "virality" in large-scale social systems.

Within the observational setting, researchers have applied quasi-experimental designs (QEDs) to perform causal inference in relational data. QEDs exploit fortuitous circumstances in data that allow for the approximation of an experimental design post-hoc. For example, Oktay et al. [20] apply QEDs to Stack Overflow, an online question and answer site for programming, to understand the dynamics of users' behavior on the site. Krishnan and Sitaraman [13] consider a quasi-experimental design to determine the relationship between network quality and user engagement with online content. Kearns et al. [10] study patterns of network formation by performing an experiment where subjects were anonymously paired, and subsequently were asked to interactively complete a graph-coloring video game.

Manski [18] considers the problem of identifiability in the potential-outcomes framework in the presence of peer influence. Ogburn and VanderWheele [19] enumerate configurations of causal graphs that result in bias from social effects on single entity, single relationship networks. Maier et al. [17] considers the more general case of *d*-separation for multi-relational domains. Maier et al. [15] apply the rules implied by Maier et al. [17] to learn the causal structure of relational domains, but explicitly do not consider inference of individual causal effects.

7. CONCLUSION AND FUTURE WORK

We have described and evaluated Relational Covariate Adjustment, an extension of nonparametric adjustment to relational data. Through the use of nonparametric regression estimators, RCA allows for estimation of a wide range of functional dependencies without modification. We showed the efficacy of this approach to causal inference with a set of experiments that examine Relational Covariate Adjustment and other experimental adjustment methods over a range of graph topologies.

This work represents one step toward a much larger goal of general causal inference in relational domains. Toward that end, we plan to extend RCA to the case of multiple entities and relationships and to extend the calculus of interventions by developing techniques that can estimate the effects of interventions that add or remove nodes or edges from the network.

References

[1] S. Aral and D. Walker. Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*, 57(9):1623–1639, 2011.

[2] P. M. Aronow and C. Samii. Estimating average causal effects under general interference. In *Summer Meeting of the Society for Political Methodology, University of North Carolina, Chapel Hill, July*, pages 19–21, 2012.

[3] E. Bakshy, D. Eckles, and M. S. Bernstein. Designing and deploying online field experiments. In *Proceedings of the 23rd international conference on World wide web*, pages 283–292. ACM, 2014.

[4] E. Bakshy, D. Eckles, R. Yan, and I. Rosenn. Social influence in social advertising: Evidence from field experiments. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 146–161. ACM, 2012.

[5] J. Bosch. Building products as innovation experiment systems. In *Software Business*, pages 27–39. Springer, 2012.

[6] D. S. Choi. Estimation of monotone treatment effects in network experiments. *arXiv preprint arXiv:1408.4102*, 2014.

[7] J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, pages 1189–1232, 2001.

[8] H. Gui, Y. Xu, A. Bhasin, and J. Han. Network a/b testing: From sampling to estimation. In *Proceedings of the 24th International Conference on World Wide Web*, pages 399–409. International World Wide Web Conferences Steering Committee, 2015.

[9] M. G. Hudgens and M. E. Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 2012.

[10] M. Kearns, S. Judd, and Y. Vorobeychik. Behavioral experiments on a network formation game. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 690–704. ACM, 2012.

[11] R. Kohavi, A. Deng, B. Frasca, T. Walker, Y. Xu, and N. Pohlmann. Online controlled experiments at large scale. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1168–1176. ACM, 2013.

[12] D. Koller. Probabilistic relational models. In *Inductive logic programming*, pages 3–13. Springer, 1999.

[13] S. S. Krishnan and R. K. Sitaraman. Video stream quality impacts viewer behavior: Inferring causality using quasi-experimental designs. *Networking, IEEE/ACM Transactions on*, 21(6):2001–2014, 2013.

[14] J. Leskovec and A. Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014.

[15] M. Maier, K. Marazopoulou, D. Arbour, and D. Jensen. A sound and complete algorithm for learning causal models from relational data. In *Uncertainty in Artificial Intelligence*, page 371. Citeseer, 2013.

[16] M. E. Maier. *Causal Discovery for Relational Domains: Representation, Reasoning, and Learning.* PhD thesis, University of Massachusetts Amherst, 2014.

[17] M. E. Maier, K. Marazopoulou, and D. D. Jensen. Reasoning about independence in probabilistic models of relational data. arXiv preprint arXiv:1302.4381, 2014.

[18] C. F. Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013.

[19] E. L. Ogburn and T. J. VanderWeele. Causal diagrams for interference. *Statist. Sci.*, 29(4):559–578, 11 2014.

[20] H. Oktay, B. J. Taylor, and D. Jensen. Causal discovery in social media using quasi-experimental designs. In *Proceedings of the SIGKDD/ACM Workshop on Social Media Analytics*, 2010.

[21] J. Pearl. Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2):226–284, 1998.

[22] J. Pearl. *Causality*. Cambridge university press, 2009.

[23] C. Perlich and F. Provost. Distribution-based aggregation for relational learning with identifier attributes. *Machine Learning*, 62(1-2):65–105, 2006.

[24] D. B. Rubin. Matching to remove bias in observational studies. *Biometrics*, pages 159–183, 1973.

[25] D. B. Rubin. Causal inference using potential outcomes. *Journal of the American Statistical Association*, 2011.

[26] P. Toulis and E. Kao. Estimation of causal peer influence effects. In *Proceedings of The 30th International Conference on Machine Learning*, pages 1489–1497, 2013.

[27] J. Ugander, B. Karrer, L. Backstrom, and J. Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–337. ACM, 2013.

8. APPENDIX

8.1 Proof of Proposition 1

PROOF. In constructing a conditional independence query with relational d-separation [15], paths composed of propositional variables, relational variables, and *intersection variables* must be considered. The set of propositional and relational variables to be considered for a perspective is directly identifiable from the relational model. Intersection variables, as defined by Maier et al. [15], are required for sound and complete reasoning of d-separation in relational domains whenever there exists two paths, $P_1 = [A, \ldots, B], P_2 = [A', \ldots, B']$ that are not subsets of each other and whose beginning and ending entity are the same, i.e., $A = A'$ and $B = B'$. We consider the case of the single entity, single relationship graph. Denote E to be the entity and R to be the relationship. Without loss of generality, we consider paths that begin at the entity. All possible path specifications then must be of the form $[A(BA)^*]$, where $*$ is the Kleene star. It follows directly that any two path specifications are either identical, or the shorter path is a sub-path of the other. This implies that for single entity, single relationship networks, intersection variables do not exist. $\square$