Doubly Robust Causal Estimation Under Multi-View Network Interference (Student Abstract)

Hanzhang Yuan[1], **Sheng Li**[1]
[1]School of Data Science, University of Virginia
{ynx9zm, shengli}@virginia.edu

Abstract

Estimating causal effects under network interference is challenging especially when edges are heterogeneous and nodes share latent dependencies. We study this realistic setting and propose **MVDR**, a targeted maximum likelihood (TMLE) framework that learns multi-view representations of covariates and exposure on heterogeneous networks while achieving *double robustness*: consistency holds if either the outcome model or the exposure density is correctly specified. MVDR supports multiple network interventions using only the observed network structure. On three semi-synthetic datasets, MVDR reduces intervention-level prediction error against baselines, and remains stable under misspecification.

Introduction

Existing causal inference methods often rely on the *Stable Unit Treatment Value Assumptions (SUTVA)* (Yao et al. 2021). However, in networked settings, interference violates SUTVA because treatment effects can spill over along edges (Chen et al. 2024; Chu, Rathbun, and Li 2021). Prior approaches either specify exposure mappings (e.g., proportion treated) or learn spillovers with graph models. However, they often assume a correctly specified outcome model and ignore heterogeneous edges and latent dependency among connected nodes. We address this gap with **MVDR**, which: (1) aggregates multi-view network information to form confounder/exposure summaries; (2) integrates TMLE to debias using the ratio between target and observed exposure densities; and (3) supports policy evaluation under static/dynamic/stochastic interventions. Experiments on three networks show consistent gains and robustness.

Problem Setting & Assumptions

Let $A^{(k)} \in \{0,1\}^{n\times n}$ be K views of network A, $X \in \mathbb{R}^{n\times d}$ covariates, $T \in \{0,1\}^n$ treatment, $Y \in \mathbb{R}^n$ outcome. Define summaries $W_i = s_X(\{X_j : A_{ij} = 1\})$ and $V_i = s_T(\{T_j : A_{ij} = 1\})$. For an intervention T^* inducing V_i^*, the target is to estimate the average potential outcome:

$$\psi_n = E[\bar{Y}_n^*] = \frac{1}{n}\sum_{i=1}^{n}\sum_{w} m(v_i^*, w) h_i^*(v_i^*.w). \quad (1)$$

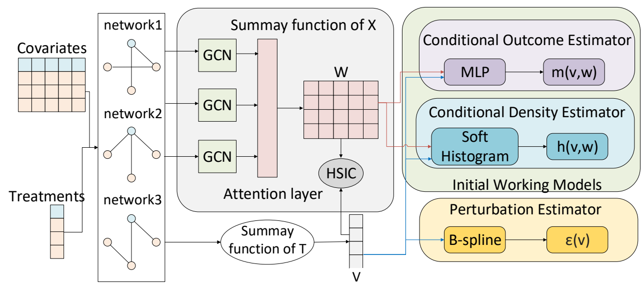

Figure 1: Model architecture of our proposed method.

with $m(v,w) = \mathbb{E}[Y \mid V = v, W = w]$ and $h(v \mid w) = P(V = v \mid W = w)$.
We assume standard positivity/exchangeability for (V, W) and a local latent dependence structure.
Assumption 1 (Local Latent Dependence Structure): The error terms $\varepsilon_{X_i} \not\perp \varepsilon_{X_j}$ (and analogously for $\varepsilon_{Y_i}, \varepsilon_{C_i}$) are permitted to be dependent if and only if nodes i and j are connected through a common neighbor, i.e., if $\exists\, k$ such that $A_{ik} = A_{kj} = 1$. The influence function $D_N(o)$ for ψ_n can be written as:

$$\begin{aligned} D_N(o) =& \frac{1}{n}\sum_{i=1}^{n}\Big(E[m(V_i^*, W_i) \mid X = x] - \psi_n + \\ & \frac{\bar{h}^*(v_i, w_i)}{\bar{h}(v_i, w_i)}\{y_i - m(v_i, w_i)\}\Big), \end{aligned} \quad (2)$$

where $\bar{h}(v_i, w_i) = \frac{1}{n}\sum_{j=1}^{n} h_j(v_i, w_i)$ and $\bar{h}^*(v_i, w_i) = \frac{1}{n}\sum_{j=1}^{n} h_j^*(v_i, w_i)$. The value of $D_n(o)$ is 0 at the true ψ_n. The influence function represents the first-order derivative of the estimand with respect to the observed data distribution, intuitively capturing how local perturbations in the outcome or exposure models affect ψ_n. It provides a correction term that balances the bias of both models, ensuring the double robustness of the estimator (Ogburn et al. 2022).

Methodology

The influence function in Eq. (2) yields a *doubly-robust* estimator of ψ_n, i.e., consistency holds if either m or h is correctly specified. Our MVDR approach aggregates the multi-view network information and then learns the two initial

Dataset	Intervention	CFR+z	ND+z	TARNet	NetEst	HINITE	TNet	Mvdr(w/o. $\mathcal{L}_3$)	Mvdr
BlogCata	Static0	1.06(0.22)	1.06(0.20)	1.05(0.32)	1.05(0.65)	1.27(0.45)	2.04(0.83)	0.87(0.40)	**0.84**(0.25)
	Static1	1.64(0.23)	1.66(0.28)	1.65(0.34)	1.64(0.46)	1.72(0.79)	27.83(9.10)	0.48(0.30)	**0.31**(0.09)
	Stochastic	1.68(0.36)	1.06(0.28)	1.70(0.52)	1.69(0.21)	1.44(0.43)	2.28(2.36)	0.81(0.30)	**0.58**(0.11)
	Dynamic	1.55(0.33)	1.66(0.34)	1.55(0.47)	1.55(0.50)	0.60(0.91)	4.18(4.21)	0.67(0.40)	**0.54**(0.32)
Flickr	Static0	0.28(0.07)	0.36(0.08)	0.16(0.07)	0.15(0.11)	0.24(0.03)	2.70(4.22)	0.75(0.16)	**0.09**(0.13)
	Static1	0.18(0.06)	0.23(0.10)	0.19(0.08)	0.19(0.06)	0.21(0.05)	2.14(4.37)	0.72(0.15)	**0.15**(0.14)
	Stochastic	0.14(0.05)	0.14(0.05)	0.26(0.03)	0.24(0.12)	0.14(0.11)	5.20(3.21)	0.15(0.08)	**0.10**(0.12)
	Dynamic	0.19(0.04)	0.23(0.07)	0.21(0.06)	0.19(0.13)	**0.10**(0.06)	46.79(48.30)	0.23(0.14)	0.12(0.08)
DBLP	Static	**0.42**(0.54)	1.23(0.38)	1.39(0.36)	1.36(0.24)	1.21(0.21)	3.64(2.40)	1.03(0.75)	0.91(0.85)
	Static1	0.54(0.21)	1.04(0.48)	1.09(0.49)	1.09(0.24)	**0.03** (0.01)	3.54(2.25)	1.34(0.79)	0.07(0.90)
	Stochastic	0.95(0.61)	1.50(0.47)	2.04(0.46)	1.50(0.39)	0.93(0.25)	3.55(1.67)	0.93(0.60)	**0.83**(1.00)
	Dynamic	1.02(0.47)	1.53(0.60)	2.36(0.65)	1.60(0.30)	1.25(0.71)	2.93(3.00)	0.78(0.57)	**0.61**(1.00)

Table 1: Comparison of methods on causal effect of network estimand under intervention: *Static0* with all individuals controlled; *Static1* with all treated; *Stochastic* with 35% randomly treated; *Dynamic* with covariate-dependent probabilities treated.

models for outcome prediction m and density estimation h. The TMLE update step ensures that the efficient influence function can be solved, so the result can be doubly-robust.

Multi-view Representations. Multiple views of a network can be constructed by capturing different relationships among nodes, such as applying various distance metrics. For each view $A^{(k)}$, a graph neural network (GCN) (Kipf and Welling 2017) produces node embeddings $g_i^{(k)}$ and attention weights α_{ik}. Attention fusion, $W_i = \sum_k \alpha_{ik} g_i^{(k)}$, represents the summary function $s_{x,i}$, which aggregates information of node i and its neighbors on the multi-view networks. Exposure summary V can be calculated as the ratio of treated neighbors in different networks. A HSIC penalty discourages redundancy between W and exposure summary V.

Initial Models. Two initial working models and one perturbation model are designed to formulate the doubly-robust estimator: (i) Outcome model $m_\theta(v, w)$ via two MLP heads (treated/control) optimized by MSE: $L_1 = \frac{1}{n}\sum_{i=1}^{n}(m_{t_i}(t_i, v_i, w_i) - y_i)^2$. (ii) For the conditional density estimator h, we first factorize it as the product of conditional density of individual treatment $h_1(t_i|w_i)$ and the neighborhood exposure $h_2(v_i|w_i)$. $h_1(t_i|w_i)$ is approximated through sigmoid activation function; the continuous exposure v is first normalized into the interval $[0, 1]$ and then scaled to a discrete grid of C bins. The density estimator is optimized with the total loss: $L_2 = \sum_{i=1}^{n}[\alpha \text{CrossEntropy}(h_1(t_i|w_i)) - \log(\hat{h_2}(v_i|w_i))]$. (iii) To achieve doubly robust estimation, we use an MLP block that learns ϵ to solve the efficient influence function. The loss function of this perturbation estimator is: $L_3 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{m}_i - \frac{\epsilon}{\hat{h}_i})^2$.

TMLE Update. For the target density h^*, we form $H_i = \frac{h^*(V_i^*, W_i)}{h(V_i, W_i)}$ and update: $\tilde{Y}_i^* = m_\theta(V_i^*, W_i) + \epsilon(V_i^*, W_i) \cdot H_i$, where $\epsilon(\cdot)$ minimizes the empirical mean of the influence function. The plug-in estimate is: $\hat{\psi}_n = \frac{1}{n}\sum_i \tilde{Y}_i^*$. Under mild regularity conditions provided, the proposed TMLE estimator remains consistent and asymptotically efficient when applied to the multi-view network setting, as it solves the efficient influence function and inherits the double robustness property from standard TMLE theory.

Experiments

Datasets. In our experiment, we utilize three semi-synthetic datasets: ***BlogCatalog***, ***Flickr*** and ***DBLP***. Since it is impossible to observe the counterfactual outcome, we mimic the unknown output following prior studies as:

$$Y_i = \beta_T \cdot T_i + f_X(X_i) + f_{NX}(X_{-i}) + f_{NT}(T_{-i}) + \varepsilon_i.$$

Evaluation Metric. We evaluate the accuracy of the model predicting the average expected outcome of a group receiving treatment under different intervention policies: $\varepsilon_{intervention} = |\widetilde{Y}_n^* - Y_n^*|$, where $Y_n^* = \frac{1}{n}\sum_{i=1}^{n} Y_i^*$.

Results and Discussions. We compare our method with several baselines, including methods with single observed network and multi-view network. CFR (Johansson et al. 2023) and ND (Guo, Li, and Liu 2019) are modified with additionally inputting the exposure. TARNet (Shalit, Johansson, and Sontag 2017) has a similar model architecture as CFR models but removes the balance term. HINITE (Lin et al. 2023) considers several observed heterogeneous networks. NetEst (Jiang and Sun 2022) is designed for effect estimation under interference; TNet (Chen et al. 2024) only considers single network and is a doubly robust method. Results in Table 1 show that MVDR outperforms baselines in most cases, especially compared with the other doubly-robust method TNet. In particular, considerable improvements are observed when the interventions are stochastic or dynamic. Compared with single-view baselines, integrating multi-view information enables the model to capture heterogeneous and complementary dependency structures across different relation types. This leads to richer representations of confounding and interference. It also improves estimation stability and accuracy under various intervention settings.

Conclusions

This paper proposes a new framework to estimate average expected outcome on an observed heterogeneous network with latent network dependency through targeted maximum likelihood estimation method. Experiments on three semi-synthetic datasets show that the proposed **MVDR** method outperforms baselines in estimating average expected outcome under hypothetical intervention policies on networks.

Acknowledgments

The work is supported in part by the U.S. Office of Naval Research Award under Grant Number N00014-24-1-2668, the National Science Foundation under Grants IIS-2316306 and CNS-2330215, and the National Institutes of Health (NIH) under Grant R01EB293388.

References

Chen, W.; Cai, R.; Yang, Z.; Qiao, J.; Yan, Y.; Li, Z.; and Hao, Z. 2024. Doubly Robust Causal Effect Estimation under Networked Interference via Targeted Learning. arXiv:2405.03342.

Chu, Z.; Rathbun, S. L.; and Li, S. 2021. Graph infomax adversarial learning for treatment effect estimation with networked observational data. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 176–184.

Guo, R.; Li, J.; and Liu, H. 2019. Learning Individual Causal Effects from Networked Observational Data. ArXiv:1906.03485 [cs].

Jiang, S.; and Sun, Y. 2022. Estimating Causal Effects on Networked Observational Data via Representation Learning. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 852–861. Atlanta GA USA: ACM. ISBN 978-1-4503-9236-5.

Johansson, F. D.; Shalit, U.; Kallus, N.; and Sontag, D. 2023. Generalization Bounds and Representation Learning for Estimation of Potential Outcomes and Causal Effects. arXiv:2001.07426.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *International Conference on Learning Representations*.

Lin, X.; Zhang, G.; Lu, X.; Bao, H.; Takeuchi, K.; and Kashima, H. 2023. Estimating Treatment Effects Under Heterogeneous Interference. ArXiv:2309.13884 [cs].

Ogburn, E. L.; Sofrygin, O.; Diaz, I.; and van der Laan, M. J. 2022. Causal inference for social network data. arXiv:1705.08527.

Shalit, U.; Johansson, F. D.; and Sontag, D. 2017. Estimating individual treatment effect: generalization bounds and algorithms. arXiv:1606.03976.

Yao, L.; Chu, Z.; Li, S.; Li, Y.; Gao, J.; and Zhang, A. 2021. A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 15(5): 1–46.