Influence Maximization for Social Network Based Substance Abuse Prevention

[1] Aida Rahmattalabi, [2] Anamika Barman Adhikari, [3] Phebe Vayanos, [3] Milind Tambe, [3] Eric Rice, [4] Robin Baker
[1] University of Southern California, rahmatta@usc.edu
SAL 300, 941 Bloom Walk, Los Angeles, California, 90089, (+1) 541-908-9553
[2] University of Denver, anamika.barmanadhikari@du.edu
[3] University of Southern California, {phebe.vayanos, tambe, ericr}@usc.edu
[4] Urban Peak Organization, Robin.Baker@urbanpeak.org

Abstract

Substance use and abuse is a significant public health problem in the United States. Group-based intervention programs offer a promising means of reducing substance abuse. While effective, inappropriate intervention groups can result in an increase in deviant behaviors among participants, a process known as *deviancy training*. In this paper, we present GUIDE, an AI-based decision aid that leverages social network information to optimize the structure of the intervention groups.

Introduction

Substance use and abuse is a significant public health problem among youth in the United States. According to the Monitoring the Future study (McCabe et al. 2014), around 54 percent of high school students have tried at least one illicit substance. Interventions programs have successfully utilized social networks to disseminate and reinforce positive behavioral norms (e.g., (Valente et al. 2003)). This is achieved through formation of subgroups where the individuals can talk, share experiences and engage in various constructive activities, and this way they form new social ties or abandon some of their existing relationships. Unfortunately, these social network-based efforts may also inadvertently increase the chances of exposure to negative social influence, as the social network of the youth changes. This is known as *deviancy training* and has been a problematic issue in these prevention programs. From an AI perspective, this problem can be viewed as a social network partitioning problem with the objective of maximizing positive influence and minimizing negative influence. However, to best of our knowledge no work has addressed such influence-based partitioning of networks with changing structures. To address this challenge, we propose an AI-based decision aid, called GUIDE (**G**ro**U**p-based **I**ntervention **DE**cision aid). GUIDE assists interventionists in substance abuse prevention, using a model for the group-based interventions that enables predicting, both the expected success of the intervention, and the possibility of harm, or deviancy training. We show that finding the optimal network partition is NP-hard and we use both a Mixed Integer Program (MIP) and a greedy-based local search method that enables us to optimize for the network partitions.

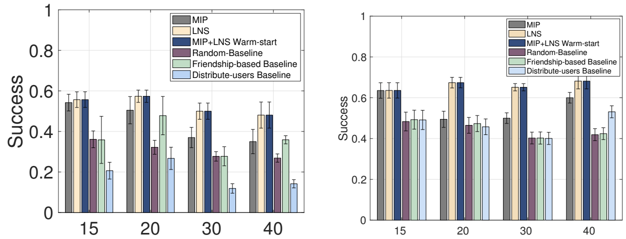
Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Same Group	no-tie	weak	strong
(user, user)	strong	strong	strong
(non-user, non-user)	strong	strong	strong
(non-user, user)	weak	weak	strong
(user, non-user)	weak	weak	strong
Separate Groups	**no-tie**	**weak**	**strong**
(user, user)	none	none	strong
(non-user, non-user)	none	weak	strong
(non-user, user)	none	none	weak
(user, non-user)	none	none	weak

Table 1: Changes in tie strength post-intervention. The existing relationships, and the behavior of the individuals as well as their assignment to groups impacts the changes.

Tie Formation and Breakage. As a result of the interventions, the strength of the relationships is subject to change (Centola and Macy 2007). For example, there is empirical evidence to suggest that the more similar two individuals are, the stronger their ties are (Aral and Walker 2014). Also, if two individuals are separated and at least one of them has "user" behavior, the intervention message will be to cut or weaken that tie. Therefore, based on behavioral theories, and observations in the previous interventions, we propose a model to explain how the network evolves during the course of the intervention which is summarized in Table 1. In this table, the row labels are the behavior of the nodes, the column labels show their pre-intervention tie and the entries indicate the post-intervention tie.

Substance Abuse Prevention Influence Spread Model. Depending on how the network evolves, we evaluate the influence to predict the changes in the nodes' behaviors. We use a variant of the popular Linear Threshold model proposed in (Borodin, Filmus, and Oren 2010). Base on our model, each node selects a threshold value, uniformly at random, to represent his/her threshold to change behavior. If the incoming signal from the opposite behavior exceeds this threshold, the change happens with a fixed probability.



(a) Synthetic Graphs (b) Real Graphs

Figure 1: Solution quality of MIP and LNS in GUIDE, against three baselines commonly employed by practitioners.

Mixed-Integer Programming Formulation

We present a Mixed Integer Linear Optimization (MIP) formulation for this problem and we use Gurobi solver to find the optimal partitioning.

Local Neighborhood Search (LNS)

We also use local search methods to optimize for the network partitions for scalability. LNS starts from a random feasible graph partition, and it improves the solution by searching in a space of candidate solutions. In this work, we restrict the search neighborhood to that created by random single swap of pairs of nodes. The search continues until no single swap can further improve the solution.

Results and Discussion

Baselines. For evaluation, we compare three variations of our optimization approach (MIP, LNS and MIP+LNS which is MIP using LNS solution as warm-start) against three different baselines that either randomly assign the individuals, or let them decide based on their friendships, or finally a teacher nominated baseline, which uses a heuristic to divide up the participants. One the common heuristics is the even distribution of the “users” across groups.

Solution Quality Metrics. Different solution strategies are compared based on a *success* metric, which we define as:

$$success = \frac{intervention\ impact}{maximum\ possible\ impact}$$

The numerator is in fact the expected number of youth that have become “non-users” as the result of the intervention. The denominator is its maximum possible value which corresponds to the case where all “users” threshold are exceeded (they are surrounded by “enough” “non-user” friends).

Solution Quality. Figures 1 (a) and (b) compare the *success* of the optimization techniques, MIP and LNS against different baselines across 4 different network sizes. The MIP solve time is given a cutoff equal to the solution time of LNS (summed over the 50 iterations). These results indicate that the solutions of both MIP+LNS and LNS are significantly better than any of the traditional methods for forming these groups, both statistically and practically. Surprisingly, in figure 1 (b) it can be observed that the common intuition of evenly distributing the “users” across the groups is in fact very sub-optimal (Distribute-users baseline). This is an insightful result for practitioners and is one of key areas where GUIDE can help practitioners. We show that uniform distribution of “users,” while ignoring their existing relationships, can greatly decrease the success rate of these interventions. The high variation in the friendship-based assignment is another important result. This is aligned with the conclusion in (Valente 2012) that network-based strategies can be very beneficial for the intervention success, while they can also cause more harm if are not carefully designed. There are computational lessons learned as well. For example, MIP is guaranteed to find the optimal solution by searching the entire solution space, but as shown here, it is not a practical solution due to time constraints. And in fact, LNS outperforms MIP solution given the same time budget. To further analyze the quality of the LNS solution, we performed a new optimization using the LNS solution as the warm up solution in MIP. These experiments are run for up to 12 hours. Interestingly, the LNS and MIP+LNS warm-start perform almost the same, providing better confidence in the LNS solution.

Conclusion

Substance abuse is a very significant public health and social problem in the United States. We showed that by careful construction of the intervention groups, we can outperform the traditional strategies significantly. GUIDE is developed in collaboration with Urban Peak, a homeless-youth serving organization in Denver, CO, and is under preparation for deployment.

References

[Aral and Walker 2014] Aral, S., and Walker, D. 2014. Tie strength, embeddedness, and social influence: A large-scale networked experiment. *Management Science* 60(6):1352–1370.

[Borodin, Filmus, and Oren 2010] Borodin, A.; Filmus, Y.; and Oren, J. 2010. Threshold models for competitive influence in social networks. In *Proceedings of the 6th international conference on Internet and network economics*, 539–550. Springer-Verlag.

[Centola and Macy 2007] Centola, D., and Macy, M. 2007. Complex contagions and the weakness of long ties. *American journal of Sociology* 113(3):702–734.

[McCabe et al. 2014] McCabe, S. E.; West, B. T.; Veliz, P.; Frank, K. A.; and Boyd, C. J. 2014. Social contexts of substance use among us high school seniors: a multicohort national study. *Journal of Adolescent Health* 55(6):842–844.

[Valente et al. 2003] Valente, T. W.; Hoffman, B. R.; Ritt-Olson, A.; Lichtman, K.; and Johnson, C. A. 2003. Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *American journal of public health* 93(11):1837–1843.

[Valente 2012] Valente, T. W. 2012. Network interventions. *Science* 337(6090):49–53.