

End to end influence maximization for HIV prevention

Bryan Wilder, Laura Onasch-Vera, Juliana Hudson, Jose Luna, Nicole Wilson, Robin Petering, Darlene Woo, Milind Tambe, Eric Rice

Center for Artificial Intelligence in Society, University of Southern California
bwilder, onaschve, jnhudson, joseluna, wilsonnj, petering, darlenew, tambe, ericr@usc.edu

Abstract. This work aims to overcome the challenges in deploying influence maximization to support community driven interventions. Influence maximization is a crucial technique used in preventative health interventions, such as HIV prevention amongst homeless youth. Drop-in centers for homeless youth train a subset of youth as peer leaders who will disseminate information about HIV through their social networks. The challenge is to find a small set of peer leaders who will have the greatest possible influence. While many algorithms have been proposed for influence maximization, none can be feasibly deployed by a service provider: existing algorithms require costly surveys of the entire social network of the youth to provide input data, and high performance computing resources to run the algorithm itself. Both requirements are crucial bottlenecks to widespread use of influence maximization in real world interventions.

To address the above challenges, this paper introduces the CHANGE agent for influence maximization. CHANGE handles the end-to-end process of influence maximization, from data collection to peer leader selection. Crucially, CHANGE only surveys a fraction of the youth to gather network data and minimizes computational cost while providing comparable performance to previously proposed algorithms. We carried out a pilot study of CHANGE in collaboration with a drop-in center serving homeless youth in a major U.S. city. CHANGE surveyed only 18% of the youth to construct its social network. However, the peer leaders it selected reached just as many youth as previously field-tested algorithms which surveyed the *entire* network. This is the first real-world study of a network sampling algorithm for influence maximization.¹

Keywords: Influence maximization · HIV prevention · social networks.

1 Introduction

This paper presents and field-tests a novel, practical agent for influence maximization, the challenge of selecting a small set of seed nodes in a social network

¹ An extended version of this paper has been accepted as a full paper at AAMAS 2018.

who will diffuse information to many others. Such techniques have important applications ranging from preventative health [13, 1] to international development [2].

We are particularly motivated by the challenge of preventing HIV spread among homeless youth [11, 18, 10] (although our contributions would also assist other public health interventions). Here, influence maximization is used to select homeless youth who will serve as *peer leaders* and spread messages about HIV prevention through their social network. Pilot studies in this domain have shown that algorithmic approaches have great promise, substantially outperforming status-quo heuristics [17]. However, current algorithms have a high barrier to entry: they require a great deal of time to gather the complete social network of the youth, expertise to select appropriate parameters, and computational power to run the algorithms. None of these are likely available to resource-strained service providers who will ultimately be the ones to deploy influence maximization.

Gathering network data is particularly onerous because it requires individually surveying upwards of a hundred youth. Further, network collection is more time intensive than simple survey methods, requiring days of time for a dedicated team of social work researchers. It is not feasible for service providers with many other responsibilities.

The other barriers are also serious impediments to wide-scale adoption of influence maximization. Service providers will not have access to the high-performance computing resources required by previous algorithms, where high computational cost is often incurred to find solutions robust to unknown parameters. For instance, DOSIM, the state of the art algorithm for robust influence maximization [15], takes hours of runtime on a high-performance computing system. In reality, a deployed system would need to run in minutes on a laptop.

This paper presents CHANGE (CompreHensive Adaptive Network samplinG for social influencE), a novel, end-to-end agent for influence maximization which addresses the above barriers via a set of algorithmic contributions. CHANGE is easy to deploy, but this simplicity is crucially enabled by a series of insights into the social structure of homeless youth (which may be useful for other vulnerable populations). We conducted a pilot test of CHANGE’s performance in a real deployment by a drop-in center serving homeless youth in a major U.S. city. CHANGE was used to plan a series of interventions designed to spread HIV awareness among the youth. *CHANGE obtained comparable influence spread to state of the art algorithms while surveying only 18% of nodes for network data.*

Overall, CHANGE offers a practical, field-tested vehicle for deployed influence maximization which drastically lowers the barrier to entry. To our knowledge, this is the first real-world pilot study of a network sampling algorithm for influence maximization and only the second ever field test of *any* influence maximization algorithm.

2 Related Work

The influence maximization problem was introduced by Kempe et al. [7], and has been extensively studied since then [3, 12, 4, 5]. Most work has focused on algorithms which are scalable to extremely large networks, primarily in the context of online viral marketing. Recently, HIV prevention (and preventative health more broadly) has emerged as a new application area for influence maximization which brings its own set of research challenges. Yadav et al. [16] proposed HEALER, a POMDP-based algorithm for selecting influential peer leaders. Subsequently, Wilder et al. [15] introduced the DOSIM algorithm which uses robust optimization to account for uncertainty about the true probability of influence propagation. Our approach to parameter robustness is similar to techniques used in robust MDP planning [8], though the domains are entirely different.

Yadav et al. [17] conducted a real-world pilot study of HEALER and DOSIM, and found that both algorithms significantly outperformed the status-quo heuristic used by agencies (selecting high-degree nodes). However, neither algorithm addresses any of the challenges described above. Both assume that the entire social network is provided as input, which is unrealistic in practice due to the enormous effort required. Further, only DOSIM handles uncertainty about the probability of influence spread, and its method for doing so is extremely computationally intensive (see Section 4.3).

Separate work by Wilder et al. [14] considered network data collection. They proposed the ARISEN algorithm which samples a portion of youth in the network to collect data from. While ARISEN can be theoretically analyzed for certain network structures, it is not practically suitable to deployment because it relies on querying a sequence of specific youth who may be difficult to locate (see Section 4.2). Moreover, ARISEN does not consider either parameter uncertainty or execution errors (the possibility that some peer leaders will not attend), both of which we incorporate into CHANGE.

3 Problem description

Motivating domain: Our work is designed to overcome the challenges in deploying influence maximization techniques to support community-driven interventions. We are specifically motivated by the challenge of raising awareness about HIV among homeless youth. Typically, an HIV awareness intervention will be provided by a drop in center or other organization which serves homeless youth. Each intervention is a day-long class followed by weekly hour-long meetings. Hence (as is typical in many intervention domains), the service provider will almost never have enough resources to deliver the intervention to all of the youth that frequent the center; instead, the intervention is usually delivered to 15-20% of the population². Further, limitations on space and personnel mean that the intervention can typically be delivered to only 4-6 youth at a given

² Note that while CHANGE directly surveys $\sim 18\%$ of youth, they name others as friends, resulting in a larger sampled graph.

time, so the training is broken up over a series of small sessions. These youth are trained as *peer leaders* who communicate with other youth about HIV prevention. This amplifies the reach of the intervention through the social network of the homeless youth. The question is which youth will make the most effective peer leaders, able to reach the greatest number of their peers. This is an *influence maximization* problem, which we now formalize.

Influence: The youth have a social network represented as a graph $G = (V, E)$. Each youth is initially inactive, meaning that they have not received information about HIV prevention. Once nodes are activated by the intervention, they have a chance to influence their peers. We model this process through a variant on the classical independent cascade model (ICM) which has been used by previous work on HIV prevention and better reflects realistic time dynamics [16, 15, 17]. The process unfolds over discrete time steps $t = 1 \dots T$, where T is a time horizon. There is a propagation probability p . When a node becomes active, it attempts to activate each of its neighbors. Each attempt succeeds independently with probability p . Activation attempts are made at each time step until either the neighbor is influenced or the time horizon is reached.

Note that the assumption that p is uniform across edges is without much loss. As noted by He and Kempe [6], a uniform p is equivalent to each edge drawing an individual propagation probability i.i.d. from a distribution with mean p . This is because the following processes are analytically equivalent: (1) propagate influence with probability p and (2) draw a propagation probability q from a distribution with $\mathbb{E}[q] = p$ and then propagate influence with probability q . Hence, our model actually subsumes *any* stochastic model where the propagation probabilities are drawn from a common prior.

Interventions: At each time step $t = 1 \dots T$, the algorithm selects a seed set A_t containing up to K nodes. However, each seed node may or may not actually attend the intervention. This problem is particularly acute with homeless youth since a number of factors could prevent a given youth from attending (e.g., being arrested, running out of money for a bus ticket, etc.). Hence, we assume that each node v has a hidden state $x_v \in \{present, absent\}$. Each node's state is drawn independently from some prior distribution D . For simplicity, we will take D to set each node to be present with probability q . However, all of our analysis applies to arbitrary distributions. For each $v \in A_t$, if $x_v = present$, then v is activated. Nothing occurs if $x_v = absent$. Note that an absent node can still become activated by others, since they may still be in contact with others in the social network. After the set A_t is chosen, the intervention occurs and the hidden state of each $v \in A_t$ is observed. We denote the set of all observations received at time t as O_t . The algorithm may use this information to plan the next intervention. In other words, the problem is *adaptive*. To model adaptivity, we introduce the notion of a *policy*. A policy maps from past actions and observations to the action that should be taken next. Let $\mathcal{A} = \{S \subseteq V : |S| \leq K\}$ be the set of all possible actions. A *history* is the current sequence of actions chosen and observations received, denoted by $\psi_t = ((A_1, O_1), (A_2, O_2), \dots, (A_t, O_t))$. Let Ψ be the set of all possible histories. A policy is a mapping $\pi : \Psi \rightarrow \mathcal{A}$. Let

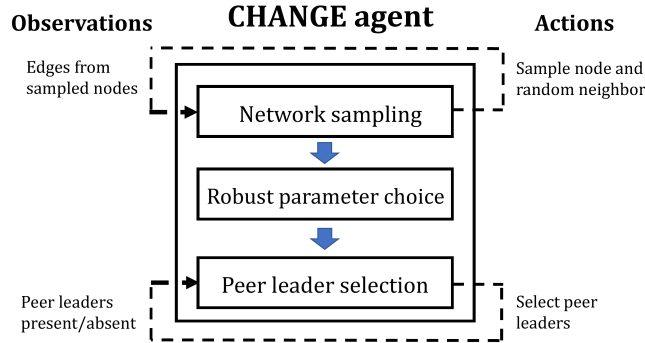


Fig. 1. Illustration of the CHANGE agent.

$A(\psi_t) = (A_1 \dots A_t)$ be the sequence of actions taken and $O(\psi) = (O_1 \dots O_t)$ be the corresponding observations (whether each peer leader was present or absent). Recall that youth are trained in groups of 4-6; the policy selects a group of youth to invite given who was trained previously. We denote the objective as $f(A(\psi)|O(\psi))$. f is the expected number of nodes influenced by the seed nodes in $A(\psi)$ conditioned on the observations in $O(\psi)$. We overload notation and let $f(\pi) = \mathbb{E}_{\psi \sim \pi}[f(A(\psi)|O(\psi))]$ be the expected reward from running policy π , where the expectation ranges over the hidden state \mathbf{x} (which determines π 's actions) as well as the influence process. Our goal is to find a policy maximizing $f(\pi)$.

Uncertainty about network structure and parameters: We consider extensions to the core adaptive influence maximization problem which account for the lack of information endemic in field deployments. First, we consider the case where the structure of the network (the edges E) are unknown. To address this challenge, we give our agent a budget of M queries to run before conducting the intervention. Each query may target either a uniformly random node, or the neighbor of a node already queried. When a node is queried, it reveals all of its edges. The goal is to use the M queries to uncover a set of edges which suffice to identify influential nodes.

We then consider an unknown propagation probability. Here, we take a robust optimization approach and look for a policy which performs well across a range of possible values for p . More detail on this part of the problem can be found in Section 4.3.

4 CHANGE: a new agent for influence maximization in the field

We now introduce the CHANGE agent for end-to-end influence maximization. Figure 1 illustrates the three components of the agent. We start with the last component, peer leader selection, since the other components exist to provide the data that the peer leader selection algorithm requires. Peer leader selection

is performed by an adaptive greedy algorithm (Algorithm 1), which handles the chance that some peer leaders may not attend the intervention and plans solutions using the observations obtained so far. Algorithm 1 requires as input a (sample of) social network and a propagation probability p . We subsequently introduce Algorithms 2 and 3 to provide these inputs.

4.1 Adaptive greedy planning

Algorithm 1 Adaptive greedy

```

1: for  $t = 1 \dots T$  do
2:    $A_t = \emptyset$ 
3:   for  $k = 1 \dots K$  do //greedily select seeds for action  $t$ 
4:      $v = \arg \max_{v \in V} \Delta(A_t \cup \{v\} | \psi_{t-1}) - \Delta(A_t | \psi_{t-1})$ 
5:      $A_t = A_t \cup \{v\}$ 
6:   execute  $A_t$  and observe  $O_t$ 
7:    $\psi_t = \psi_{t-1} + (A_t, O_t)$  //add action/observation to history

```

Given as input the graph G and propagation probability p , finding the optimal policy is a difficult planning problem. There are 2^n possible hidden states and $\binom{n}{K}$ possible actions. While it is possible to formulate the problem as a POMDP, these exponentially large state and action spaces place even small instances beyond the reach of off-the-self solvers. Hence, we exploit the structure of the problem to formulate a scalable greedy algorithm which obtains (provably) near-optimal solutions.

Pseudocode for adaptive greedy, our online planning algorithm, can be found in Algorithm 1. Algorithm 1 selects the action at each step which maximizes the expected gain in influence spread, conditioned on the observations received so far. Then, it waits until this action has been executed, observes which peer leaders attended the intervention, and greedily plans the next step. Formally, let $\Delta(A_t | \psi_{t-1}) = f(A(\psi_{t-1}) \cup A_t | O(\psi_{t-1})) - f(A(\psi_{t-1}) | O(\psi_{t-1}))$ denote the expected marginal gain to selecting A_t at time t . The greedy policy is to select $A_t = \arg \max_{|A| \leq K} \Delta(A | \psi_{t-1})$ (the outer loop of Algorithm 1). However, computing the maximizing action is itself computationally intractable (as there are $\binom{n}{K}$ possible choices). Hence, Algorithm 1 uses an additional greedy inner loop which greedily selects the elements of A_t one at a time (lines 3-5). Note that Δ can be computed by averaging over random simulations over both the hidden state (which nodes are present/absent) as well as how influence spreads via the ICM.

We prove the following theorem, which shows that greedy planning is sufficient to obtain a guaranteed approximation ratio:

Theorem 1. *Let π_G be Algorithm 1's greedy policy and π_* be an optimal policy. It holds that $f(\pi_G) \geq \left(\frac{e-1}{2e-1}\right) f(\pi_*)$.*

A proof may be found in the supplemental material. We use the *adaptive submodularity* framework of Golovin and Krause, which generalizes the classical notion of a submodular set function to adaptive policies. Their framework does not straightforwardly apply to our problem since our algorithm selects a *sequence* of actions, not a set. The order in which actions are selected matters since peer leaders who are selected earlier will have more time to influence others. We show that our problem can be reformulated as maximizing an adaptive submodular set function subject to a more complex set of constraints (a partition matroid). *This is the first approximation guarantee for adaptive influence maximization under execution errors, which is a well-known challenge in domains such as ours [15, 17].*

4.2 Network collection

Algorithm 2 Network sampling

```

1: input: vertex set  $V$ , budget  $M$ 
2:  $E = \emptyset$  //set of edges observed
3:  $S = \emptyset$  //set of nodes surveyed
4: for  $i = 1 \dots \frac{M}{2}$  do
5:   Sample  $v$  uniformly at random from  $V \setminus S$ 
6:    $S = S \cup \{v\}$ 
7:    $E = E \cup \{(v, u) : u \in N(v)\}$ 
8:   Sample  $u$  uniformly at random from  $N(v) \setminus S$ 
9:    $E = E \cup \{(u, w) : w \in N(u)\}$ 
10:   $S = S \cup \{u\}$ 
11: return  $E$ 

```

The adaptive greedy algorithm assumes that the graph G is fully specified. However, in order for an intervention to be deployed in practice, the social network needs to be laboriously gathered by interviewing the entire population of homeless youth (potentially hundreds of youth in total). This is not practical for a service provider to carry out on their own. We present an approach (Algorithm 2) which randomly samples a small number of youth to survey. Our procedure is easy for a service provider to implement in the field without much computational assistance. This simplicity is enabled by underlying insights about the structure of homeless youth social networks, which may assist with intervention design in other vulnerable populations.

We assume that the service provider has the ability to survey up to M youth. Each youth, when surveyed, reveals all of their edges. Algorithm 2 chooses $\frac{M}{2}$ nodes uniformly at random from the population to survey (line 5). For each surveyed node, it chooses a uniformly random neighbor to survey as well (line 8). Lastly, it returns the graph consisting of the reported edges. The intuition for why this procedure succeeds is that it leverages the *friendship paradox*: a phenomena where a random node's neighbor has more friends, on average, than the

node itself. Essentially, high-degree nodes are overrepresented when we sample a random neighbor instead of a uniformly random node. Thus, Algorithm 2 is disproportionately likely to find central nodes in the network who will reveal many edges and may be good potential seeds.

4.3 Parameter robustness

Algorithm 3 Robust parameter selection

```

1: input: parameter values  $p_1 \dots p_L$ 
2: for  $i = 1 \dots L$  do
3:   for  $j = 1 \dots L$  do
4:      $g(p_i, p_j) =$  value obtained by Algorithm 1 using  $p_i$  evaluated under  $p_j$ 
5: return  $\arg \max_{i=1 \dots L} \min_{j=1 \dots L} \frac{g(p_i, p_j)}{g(p_j, p_j)}$ 

```

A further complication is that the adaptive greedy algorithm assumes that the propagation probability p is known, in order to calculate the marginal gain Δ . However, p is never known precisely in practice; each intervention takes months to deploy so we are unlikely to observe the many repeated cascades needed to for learning-based approaches. Previous work has attempted to resolve this dilemma via *robust influence maximization* [6, 3, 9, 15] which finds a seed set which performs well in the worst case over an uncertainty set of possible parameters. However, the only previous work which addresses robust influence maximization in an adaptive domain is the DOSIM algorithm. DOSIM requires hours or even days of runtime on a high-performance computing cluster because it needs to brute force over a grid of possible parameter settings. Such computational expense is far beyond the capabilities of the average service provider, motivating the development of lightweight but effective heuristics for robust influence maximization.

Algorithm 3 gives the heuristic used by CHANGE. It searches for a good nominal value of the parameter p , which (when given to Algorithm 1) will result in high performance no matter what the true value of p actually is. We first discretize the interval $[0, 1]$ into L points $p_1 \dots p_L$. Let $g(p_i, p_j)$ denote the expected influence obtained when we run adaptive greedy planning based on propagation probability p_i , but the true parameter is p_j . We then find $p^* = \arg \max_{i=1 \dots L} \min_{j=1 \dots L} \frac{g(p_i, p_j)}{g(p_j, p_j)}$. Here, $\frac{g(p_i, p_j)}{g(p_j, p_j)}$, is the ratio of the value based on planning with parameter p_i to the value that could have been obtained if we knew the true parameter p_j . p^* is the parameter which maximizes the worst-case value of this ratio. Notably, this procedure requires only L^2 runs of adaptive greedy; we take $L = 10$ in practice. By contrast, DOSIM requires $O\left(\frac{n}{\epsilon}\right)^3$ runs of a greedy algorithm to achieve approximation error ϵ . This quickly reaches thousands (or tens of thousands) of runs even for moderately sized networks and requires high-performance computing resources.

5 Pilot study procedure

Table 1. Number of youth recruited, trained, and retained for follow-up in each study. CHANGE refers to the study conducted in this work to test the CHANGE agent. The other columns are taken from Yadav et al. [17], who conducted pilot tests of HEALER and DOSIM.

	CHANGE	HEALER	DOSIM
Youth recruited	64	62	56
Queried for links	18.75%	100%	100%
PL trained	15.6%	17.7%	17.85%
Retained	54.7%	73%	73%

The major contribution of this work is carrying out a pilot study which tests the CHANGE agent in a field deployment at a real drop in center serving homeless youth in a major U.S. city. Here, we outline the procedure followed for the pilot study. There were two studies, the feasibility study and the intervention study. In the feasibility study, we just tested the first component of CHANGE (network data collection) to validate that it works in practice to gather high-quality data. In the intervention study, we carried out actual interventions with homeless youth at the center. This step used all three steps of the CHANGE agent: we gathered the network, found a robust set of parameters, and then carried out interventions.

For each of the studies, we enrolled (respectively) 72 and 64 youth. Each youth was paid \$20 to enroll in the study (all monetary incentives were the same as prior studies [17]). We ran CHANGE’s data collection mechanism, randomly sampling a subset of youth to query for ties. Each youth who enrolled was also asked to complete a baseline survey. As part of this survey, we also gathered the *full* network consisting of ties from all of the youth. *We emphasize that this data was collected just for analysis. We did not use the full network to plan interventions, and we would not expect an agency to conduct this step in a regular deployment.*

In the intervention study, trained social workers delivered the *Have You Heard* intervention, previously published in the public health literature [11]. The social workers conducted a day-long class with the selected youth, covering HIV awareness and prevention, and training the youth as peer leaders to communicate with other youth at the agency. Peer leaders were paid \$60. Three sets of peer leaders were selected by CHANGE, with approximately 4 peer leaders in each set. This matches the size of trainings used in previous influence maximization pilot studies [17]. Table 1 reports specific values on the number of youth enrolled, queried for edges, and trained as peer leaders for our pilot test as well as pilot tests of previous algorithms. One month after the start of the study, we conducted a follow up survey with all of the youth who initially enrolled. Some

youth were lost to follow up (see Table 1). We asked the youth whether they had received information about HIV prevention from a peer who was part of the study. Youth were paid \$20 to respond to the follow up survey. We emphasize that all aspects of the intervention study (the training materials for peer leaders, survey instruments, etc.) are identical to Yadav et al. [17], so our results are directly comparable.

6 Pilot study results

We focus here on results from the intervention study, which tested the entirety of the CHANGE agent. In this study, we recruited a population of 64 homeless youth from a drop-in center. Table 1 gives the total number of youth recruited for different activities, as well as the corresponding figures for previous pilot tests of the HEALER and DOSIM algorithms by Yadav et al. [17]. We gathered the full social network from all 64 youth, and in parallel ran Algorithm 2 with a budget of $M = 12$ youth to collect a sampled network (querying 18.75% of youth in total for links). *Only the sampled network was used to plan interventions; the full network was gathered only for analysis.* We then ran the CHANGE policy for three steps, training 10 total peer leaders (15.6% of the network). This percentage is comparable to previous studies (HEALER and DOSIM trained approximately 17% of the network each). However, HEALER and DOSIM used the entire network to plan their intervention, compared to the 18.75% of sampled youth used by CHANGE. At one month, we conducted a follow-up survey to assess whether youth received information about HIV prevention from the peer leaders. 54.7% of youth were retained in the follow-up survey, which is a somewhat lower percentage than in previous studies. Nevertheless, we obtain a population of 34 youth who provided follow-up data.

6.1 Influence spread results

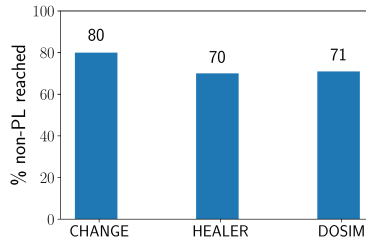


Fig. 2. Percentage of youth who were not peer leaders reached by each algorithm in its respective real-world pilot test.

We now present our core result: the number of youth who received a message about HIV prevention. We examine the percentage of youth in the follow-up

group who were not peer leaders (and hence eligible to become influenced) who reported receiving information. Figure 2 shows this percentage for our pilot study of CHANGE as well as the percentages reported by Yadav et al. [17] in their pilot studies of the state of the art algorithms HEALER and DOSIM. CHANGE reached 80% of non-peer leaders compared to approximately 70% for each of HEALER and DOSIM. *Thus, CHANGE was able to reach just as many youth while gathering data from only 18.75% of the network.* The 10% difference between CHANGE and HEALER/DOSIM could be attributable to random variation; we do not claim that CHANGE is actually more effective than algorithms which gather the entire network. Nevertheless, this result provides empirical evidence that CHANGE can perform comparably to existing state of the art influence maximization agents while drastically reducing the amount of data required.

7 Discussion and conclusion

This paper presents the CHANGE agent for influence maximization, a multi-agent problem with many applications in preventative health and other domains. CHANGE addresses major barriers to the deployment of influence maximization by service providers through a series of algorithmic contributions, backed by simulation results on real-world networks. We then conducted a real-world pilot study of CHANGE with a drop-in center serving homeless youth, the first such pilot study of sampling-based influence maximization and only the second study testing *any* influence maximization agent in the real world. CHANGE obtained comparable influence spread to previously field tested algorithms, but surveyed only 18% of youth to obtain network data. CHANGE has empirical promise in delivering high-quality influence maximization solutions in a manner which can be feasibly implemented by a service provider.

While the algorithms underlying CHANGE are easy to implement, they draw on a series of insights into the social behavior of homeless youth. One lesson learned from our study is that, to be successful in the field, algorithms must be designed with their target population and setting in mind. CHANGE both navigates challenges specific to homeless youth (e.g., the difficulty of locating youth to query for edges or serve as peer leaders) and leverages properties of their social network (the friendship paradox). Our experience shows that accounting for both challenges and opportunities in the target population is crucial to produce a practically deployable algorithm.

Acknowledgments: This research was supported by MURI Grant W911NF-11-1-0332, the California HIV/AIDS Research Program, and a NSF Graduate Research Fellowship.

References

1. Alexander, C., Piazza, M., Mekos, D., Valente, T.: Peers, schools, and adolescent cigarette smoking. *Journal of adolescent health* **29**(1), 22–30 (2001)

2. Banerjee, A., Chandrasekhar, A.G., Duflo, E., Jackson, M.O.: The diffusion of microfinance. *Science* **341**(6144), 1236498 (2013)
3. Chen, W., Wang, C., Wang, Y.: Scalable influence maximization for prevalent viral marketing in large-scale social networks. In: *KDD*. pp. 1029–1038. ACM (2010)
4. Cohen, E., Delling, D., Pajor, T., Werneck, R.F.: Sketch-based influence maximization and computation: Scaling up with guarantees. In: *CIKM* (2014)
5. Goyal, A., Lu, W., Lakshmanan, L.V.: Celf++: optimizing the greedy algorithm for influence maximization in social networks. In: *WWW* (2011)
6. He, X., Kempe, D.: Robust influence maximization. In: *KDD* (2016)
7. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. In: *KDD*. pp. 137–146. ACM (2003)
8. Kwak, J.y., Varakantham, P., Maheswaran, R., Tambe, M., Hayes, T., Wood, W., Becerik-Gerber, B.: Towards robust multi-objective optimization under model uncertainty for energy conservation. In: *AAMAS ATES workshop* (2012)
9. Lowalekar, M., Varakantham, P., Kumar, A.: Robust influence maximization. In: *AAMAS*. pp. 1395–1396 (2016)
10. Rice, E., Milburn, N.G., Rotheram-Borus, M.J.: Pro-social and problematic social network influences on hiv/aids risk behaviours among newly homeless youth in los angeles. *AIDS care* **19**(5), 697–704 (2007)
11. Rice, E., Tulbert, E., Cederbaum, J., Barman Adhikari, A., Milburn, N.G.: Mobilizing homeless youth for hiv prevention: a social network analysis of the acceptability of a face-to-face and online social networking intervention. *Health education research* **27**(2), 226–236 (2012)
12. Tang, Y., Xiao, X., Shi, Y.: Influence maximization: Near-optimal time complexity meets practical efficiency. In: *SIGMOD* (2014)
13. Valente, T.W., Pumpuang, P.: Identifying opinion leaders to promote behavior change. *Health Education & Behavior* (2007)
14. Wilder, B., Immorlica, N., Rice, E., Tambe, M.: Maximizing influence in an unknown social network. In: *AAAI Conference on Artificial Intelligence* (2018)
15. Wilder, B., Yadav, A., Immorlica, N., Rice, E., Tambe, M.: Uncharted but not uninfluenced: Influence maximization with an uncertain network. In: *AAMAS*. pp. 1305–1313 (2017)
16. Yadav, A., Chan, H., Xin Jiang, A., Xu, H., Rice, E., Tambe, M.: Using social networks to aid homeless shelters: Dynamic influence maximization under uncertainty. In: *AAMAS*. pp. 740–748 (2016)
17. Yadav, A., Wilder, B., Rice, E., Petering, R., Craddock, J., Yoshioka-Maxwell, A., Hemler, M., Onasch-Vera, L., Tambe, M., Woo, D.: Influence maximization in the field: The arduous journey from emerging to deployed application. In: *AAMAS*. pp. 150–158 (2017)
18. Young, S.D., Rice, E.: Online social networking technologies, hiv knowledge, and sexual risk and testing behaviors among homeless youth. *AIDS and Behavior* **15**(2), 253–260 (2011)