POMDPs for Assisting Homeless Shelters - Computational and Deployment Challenges

Amulya Yadav, Hau Chan[1], Albert Jiang[1], Eric Rice, Ece Kamar[2], Barbara Grosz[3], Milind Tambe
University of Southern California, CA 90089 [1]Trinity University, TX 78212
[2]Microsoft Research, WA 98052 [3]Harvard University, MA 02138
{amulyaya, ericr, tambe}@usc.edu [1]{hchan, xjiang}@trinity.edu
[2]eckamar@microsoft.com [3]grosz@eecs.harvard.edu

ABSTRACT

This paper looks at challenges faced during the ongoing deployment of HEALER, a POMDP based software agent that recommends sequential intervention plans for use by homeless shelters, who organize these interventions to raise awareness about HIV among homeless youth. HEALER's sequential plans (built using knowledge of social networks of homeless youth) choose intervention participants strategically to maximize influence spread, while reasoning about uncertainties in the network. In order to compute its plans, HEALER (i) casts this influence maximization problem as a POMDP and solves it using a novel planner which scales up to previously unsolvable real-world sizes; (ii) and constructs social networks of homeless youth at low cost, using a Facebook application. HEALER is currently being deployed in the real world in collaboration with a homeless shelter. Initial feedback from the shelter officials has been positive but they were surprised by the solutions generated by HEALER as these solutions are very counterintuitive. Therefore, there is a need to justify HEALER's solutions in a way that mirrors the officials' intuition. In this paper, we report on progress made towards HEALER's deployment and detail first steps taken to tackle the issue of explaining HEALER's solutions.

1. INTRODUCTION

HIV is a huge problem among homeless youth. Past statistics show that homeless youth are 10X more likely to get infected by HIV compared to stably housed youth [4]. The primary reason behind this is that homeless youth tend to engage in high HIV risk behaviors such as unprotected sex, sharing needles while using drugs, etc., due to an absence of educated parental figures in their life who can advise them against such high-risk activities.

Often, homeless youth do not have access to traditional health care facilities, which makes early detection, treatment and control of HIV especially challenging among homeless youth populations. To that end, many homeless shelters provide free HIV testing clinics for homeless youth to promote a habit of getting regular HIV tests among youth. Despite these facilities, homeless youth do not get tested regularly as most of them are not aware of basic information about how HIV spreads and how can it be treated. Therefore, getting regular HIV tests is not a pressing concern for them as they are not aware of the consequences of HIV infection. Thus, there is an urgent need to raise awareness about basic HIV related information among homeless youth.

To address this need, many homeless shelters conduct "intervention camps" among homeless youth to raise general awareness about HIV. These intervention camps consist of day/week long educational sessions in which youth are provided with resources and information about HIV prevention and treatment measures [21]. For example, they are provided emergency contact numbers of newly opened HIV testing centers. Free contraceptives are also distributed among them. However, financial and manpower constraints faced by homeless shelters means that they can only organize a limited number of intervention camps. Moreover, in each camp, the shelters can only manage small groups of youth ($\sim$3-4) at a time (as emotional and behavioral problems of youth makes management of bigger groups difficult). Thus, the shelters prefer a series of small sized camps organized sequentially [20]. Using these interventions, the shelter plans to maximize the spread of awareness (about HIV) among the target population (via word-of-mouth influence). To achieve this goal, the shelter uses the friendship based social network of the target population to strategically choose the participants of their limited intervention camps. Unfortunately, the shelters' job is further complicated by a lack of complete knowledge about the social network's structure [18]. Some friendships in the network are known with certainty whereas there is uncertainty about other friendships.

Thus, the shelters face an important challenge: they need a sequential plan to choose the participants of their sequentially organized interventions. This plan must address four key points: (i) it must deal with network structure uncertainty; (ii) it needs to take into account new information uncovered during the interventions, which reduces the uncertainty in our understanding of the network; (iii) the plan needs to be deviation tolerant, as sometimes homeless youth may refuse to be an intervention participant, thereby forcing the shelter to modify its plan; (iv) the intervention approach should address the challenge of gathering information about social networks of homeless youth, which usually costs thousands of dollars and many months of time [20].

In previous work, the authors presented HEALER [29], an adaptive software agent for solving this problem faced by homeless shelters. HEALER casts this problem as a Partially Observable Markov Decision Process (POMDP) and solves it using HEAL, a novel POMDP planner which quickly generates high-quality recommendations (of intervention participants) for homeless shelter officials. Our results from the previous paper show that HEALER significantly outperforms state-of-the-art techniques in terms of influence spread achieved.

HEALER is currently being deployed in a real-world pilot study,

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

in collaboration with Safe Place for Youth[1], a homeless shelter which provides food and lodging to homeless youth aged 12-25. They provide these facilities for ∼55-60 homeless youth every day. They also operate an on-site medical clinic where free HIV and Hepatitis-C testing is provided. HEALER was reviewed by officials at Safe Place for Youth and their feedback has mostly been positive. In this paper, we report on preliminary progress made in the deployment of HEALER in our pilot study.

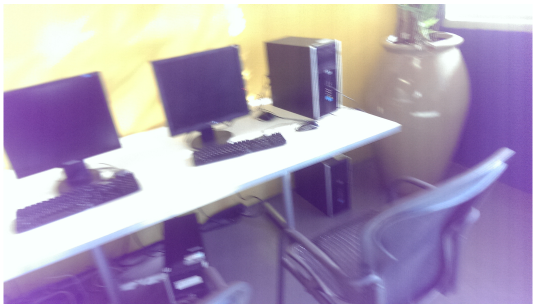


Figure 1: Computers at Safe Place for Youth where HEALER is deployed

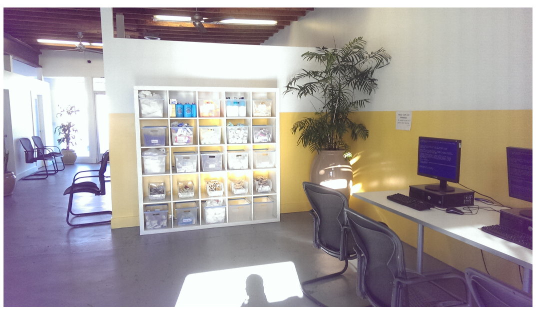


Figure 2: Emergency Resource Shelf at Safe Place for Youth

Figure 3: Enrolling homeless youth into pilot study at Safe Place for Youth

However, despite the shelter officials liking HEALER, and allowing us to conduct this pilot study with their youth, the shelter officials were surprised by HEALER's solutions. This is because HEALER's solutions maximize expected utilities (as explained later), while the homeless shelter officials pick youth for interventions based on their popularity in the network. To that end, we aim to develop a POMDP *explanation system*, which will justify HEALER's solutions to the officials in an intuitive manner. In this paper, we explain first steps taken towards building this POMDP explanation system.

[1]http://safeplaceforyouth.org

2. RELATED WORK

First, we discuss work related to influence maximization. There are many algorithms for finding 'seed sets' of nodes to maximize influence spread in networks [9, 13, 1, 27]. However, all these algorithms assume *no uncertainty in the network structure* and select a single seed set. In contrast, HEALER selects several seed sets sequentially to select intervention participants. Also, HEALER takes into account uncertainty about the network structure and influence status of network nodes (i.e., whether a node is influenced or not). Finally, unlike [9, 13, 1, 27], we use a different diffusion model as we explain later. Golovin et. al. [7] introduced adaptive submodularity and discussed adaptive sequential selection (similar to our problem), and they proved that a Greedy algorithm has a $(1 - 1/e)$ approximation guarantee. However, unlike our work, they assume no uncertainty in network structure. Also, while our problem can be cast into the adaptive stochastic optimization framework of [7], our influence function is not adaptive submodular, because of which their Greedy algorithm loses its approximation guarantees [29]. This loss of adaptive submodularity is partly due to the feedback structure that our real world domain imposes on us.

Next, we discuss literature from *social work*. The general approach to these interventions is to use Peer Change Agents (PCA) (i.e., peers who bring about change in attitudes) to engage homeless youth in interventions, but most studies don't use network characteristics to choose these PCAs [23]. A notable exception is Valente et. al. [28], who proposed selecting intervention participants with highest *degree centrality* (the most ties to other homeless youth). However, previous studies [2, 30] show that *degree centrality* performs poorly, as it does not account for potential overlaps in influence of two high degree centrality nodes.

Another field of related work is planning for reward/cost optimization. In POMDP literature, a lot of work has been done on offline planning; some notable offline planners include GAPMIN [16] and Symbolic Perseus [26]. However, since online planners scale up much better [15],we only focus on the literature on Monte-Carlo (MC) sampling based online POMDP solvers since this approach allows significant scale-up [22]. A recent paper [30] introduced PSINET-W, a MC sampling based online POMDP planner. As we show later, HEALER scales up whereas PSINET fails to do so. HEALER's algorithmic approach also offers significant novelties in comparison with PSINET. A recent paper [14] looks at the case that not all nodes in the network are known ahead of time (as opposed to our work where we only assume that some edges are not known ahead of time). However, unlike HEALER, they do not consider sequential selection of node subsets.

The final field of related work is on explanation systems for POMDPs. Khan et. al. [10] came up with template based explanation system and introduced the notions of minimal sufficient explanations. Also, Seegebarth et. al. [24] presented the hybrid plan explanation framework. However, most of these approaches deal with fully observable Markov Decision Processes. We plan on building on the ideas in these papers to build our POMDP explanation system. Next, we will give a brief overview of HEALER's design.

3. HEALER'S DESIGN

We now explain the high-level design of HEALER. It consists of two major components: (i) a Facebook application for gathering information about social networks; and (ii) a DIME Solver, which solves the DIME problem (introduced in Section 5). We first explain HEALER's components and then explain HEALER's design.

Facebook Application: HEALER gathers information about social ties in the homeless youth social network by interacting with youth via a Facebook application. We choose Facebook for gathering information as Young et. al. [32] show that ∼80% of homeless youth are active on Facebook. Once a fixed number of homeless youth register in the Facebook application, HEALER parses the Facebook contact lists of all the registered homeless youth and generates the social network between these youth. HEALER adds a link between two people, if and only if both people are (i) friends on Facebook; and (ii) are registered in its Facebook application. Unfortunately, there is *uncertainty* in the generated network as friendship links between people who are only friends in real-life (and not on Facebook) are not captured by HEALER.

Thus, HEALER's Facebook application assists homeless shelters in quickly generating first approximations of these social ties at virtually no cost. Previously, homeless shelters gathered this social network information via tedious face-to-face interviews with homeless youth, a process which cost thousands of dollars and many months of time. HEALER's Facebook application allows homeless shelters to quickly generate a (partial) homeless youth social network at low cost. This Facebook application has been tested rigorously by our collaborating homeless shelter with positive feedback and in this paper, we present some initial results using this Facebook application (see Section 9).

DIME Solver: The DIME Solver then takes the approximate social network (generated by HEALER's Facebook application) as input and solves the DIME problem (formally defined in Section 5) using HEAL [29]. HEALER provides the solution of this DIME problem as a series of recommendations (of intervention participants) to homeless shelter officials.

HEALER Design: HEALER's design (shown in Figure 4), begins with the Facebook application constructing an *uncertain* network (as explained above). HEALER has a *sense-reason-act* cycle; where it repeats the following process for T interventions.

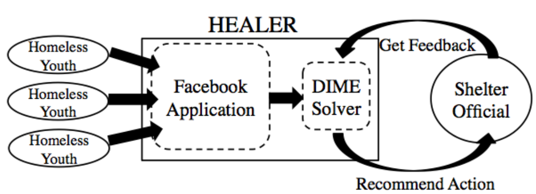


Figure 4: HEALER's Design

It *reasons* about different long-term plans to solve the DIME problem, it *acts* by providing DIME's solution as a recommendation (of intervention participants) to homeless shelter officials. The officials may choose to not use HEALER's recommendation in selecting their intervention's participants. Upon the intervention's completion, HEALER *senses* feedback about the conducted intervention from the officials. This feedback includes new observations about the network, e.g., uncertainties in some links may be resolved as intervention participants are interviewed by the shelter officials (explained more in Section 5). HEALER uses this feedback to update and improve its future recommendations.

4. NETWORK GENERATION

First, we explain our model for influence spread in *uncertain social networks*. Then, we describe how HEALER generates a social network using its' Facebook application.

4.1 Background

We represent social networks as directed graphs (consisting of *nodes* and *directed edges*) where each *node* represents a person in the social network and a *directed edge* between two nodes A and B (say) represents that node A *considers* node B as his/her friend. *We assume directed-ness of edges as sometimes homeless shelters assess that the influence in a friendship is very much uni-directional; and to account for uni-directional follower links on Facebook.* Otherwise friendships are encoded as two uni-directional links. Further, even in the case of a bi-directional friendship, the influence propagation is not symmetric in either direction of the edge and we account for this by maintaining two uni-directional links (each with a different propagation probability) for each bi-directional link.

Uncertain Network: The uncertain network is a directed graph $G = (V, E)$ with $|V| = N$ nodes and $|E| = M$ edges. The edge set E consists of two disjoint subsets of edges: E_c (the set of certain edges, i.e., friendships which we are certain about) and E_u (the set of uncertain edges, i.e., friendships which we are uncertain about). Note that uncertainties about friendships exist because HEALER's Facebook application misses out on some links between people who are friends in real life, but not on Facebook.

To model the uncertainty about missing edges, every uncertain edge $e \in E_u$ has an existence probability $u(e)$ associated with it, which represents the likelihood of "existence" of that uncertain edge. For example, if there is an uncertain edge (A, B) (i.e., we are unsure whether node B is node A's friend), then $u(A, B) = 0.75$ implies that B is A's friend with a 0.75 chance. In addition, each edge $e \in E$ (both certain and uncertain) has a propagation probability $p(e)$ associated with it. A propagation probability of 0.5 on directed edge (A, B) denotes that if node A is influenced (i.e., has information about HIV prevention), it influences node B (i.e., gives information to node B) with a 0.5 probability in each subsequent time step (our full influence model is defined below). This graph G with all relevant $p(e)$ and $u(e)$ values represents an uncertain network and serves as an input to the DIME problem. Figure 5 shows an uncertain network on 6 nodes (A to F) and 7 edges. The dashed and solid edges represent uncertain (edge numbers 1, 4, 5 and 7) and certain (edge numbers 2, 3 and 6) edges, respectively. Next, we explain the influence diffusion model that we use in HEALER.

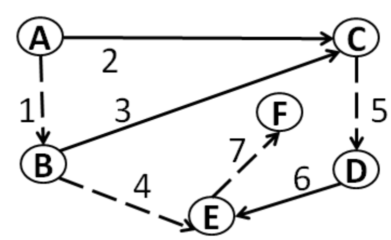


Figure 5: Uncertain Network

Influence Model We use a variant of the independent cascade model [31]. In the standard independent cascade model, all nodes that get influenced at round t get a **single** chance to influence their un-influenced neighbors at time $t+1$. If they fail to spread influence in this **single** chance, they don't spread influence to their neighbors in future rounds. Our model is different in that we assume that nodes get **multiple** chances to influence their un-influenced neighbors. If they succeed in influencing a neighbor at a given time step t', they stop influencing that neighbor for all future time steps. Otherwise, if they fail in step t', they try to influence again in the next round. This variant of independent cascade has been shown to empirically provide a better approximation to real influence spread than the standard independent cascade model [3, 31]. Further, we assume that nodes that get influenced at a certain time step remain influenced for all future time steps. We now explain how HEALER generates an *uncertain social network*.

4.2 HEALER's Facebook application

HEALER generates an *uncertain network* by (i) using its Facebook application to generate a network with no uncertain edges; (ii) using well known link prediction techniques such as KronEM [11] to infer existence probabilities $u(e)$ for all possible *missing* edges that are not present in the network; (iii) deciding on a threshold probability τ (in consultation with homeless shelter officials), so that we *only* add a *missing* edge as an uncertain edge if its inferred existence probability $u(e) > \tau$; and (iv) asking homeless shelter officials to provide $p(e)$ estimates for network edges.

Choosing τ: Rice et. al [19] show that real-world homeless youth networks are *relatively sparse*. Thus, shelter officials choose the threshold probability value τ such that the number of uncertain edges that get added because of τ does not make our input uncertain network *overly dense*. Next, we introduce the DIME problem.

5. DIME PROBLEM

We now provide some background information that helps us define a precise problem statement for DIME. After that, we will show some hardness results about this problem statement.

Given the *uncertain network* as input, HEALER runs for T *rounds* (corresponding to the number of interventions organized by the homeless shelter). In each round, HEALER chooses K *nodes* (youth) as intervention participants. These participants are assumed to be influenced post-intervention with certainty. Upon influencing the chosen nodes, HEALER '*observes*' the true state of the *uncertain edges* (friendships) out-going from the selected nodes. This translates to asking intervention participants about their 1-hop social circles, which is within the homeless shelter's capabilities [19].

After each round, influence spreads in the network according to our influence model for L *time steps*, before we begin the next round. This L represents the time duration in between two successive intervention camps. *In between rounds, HEALER does not observe the nodes that get influenced during L time steps.* HEALER only knows that explicitly chosen nodes (our intervention participants in all past rounds) are influenced. Informally then, given an uncertain network $G_0 = (V, E)$ and integers T, K, and L (as defined above), HEALER finds an online policy for choosing *exactly* K nodes for T successive rounds (interventions) which maximizes influence spread in the network at the end of T rounds.

We now provide notation for defining HEALER's policy formally. Let $\boldsymbol{\mathcal{A}} = \{A \subset V \text{ s.t. } |A| = K\}$ denote the set of K sized subsets of V, which represents the set of possible choices that HEALER can make at every time step $t \in [1, T]$. Let $A_i \in \boldsymbol{\mathcal{A}}\, \forall i \in [1, T]$ denote HEALER's choice in the i^{th} time step. Upon making choice A_i, HEALER '*observes*' uncertain edges adjacent to nodes in A_i, which updates its understanding of the network. Let $G_i\ \forall\, i \in [1, T]$ denote the uncertain network resulting from G_{i-1} with *observed* (additional edge) information from A_i. Formally, we define a history $H_i\ \forall\, i \in [1, T]$ of length i as a tuple of past choices and observations $H_i = \langle G_0, A_1, G_1, A_2, .., A_{i-1}, G_i\rangle$. Denote by $\boldsymbol{\mathcal{H}_i} = \{H_k \text{ s.t. } k \leqslant i\}$ the set of all possible histories of length less than or equal to i. Finally, we define an i-step policy $\boldsymbol{\Pi_i} \colon \boldsymbol{\mathcal{H}_i} \to \boldsymbol{\mathcal{A}}$ as a function that takes in histories of length less than or equal to i and outputs a K node choice for the current time step. We now provide an explicit problem statement for DIME.

PROBLEM 1. ***DIME Problem*** *Given as input an uncertain network $G_0 = (V, E)$ and integers T, K, and L (as defined above). Denote by $\mathcal{R}(H_T, A_T)$ the expected total number of influenced nodes at the end of round T, given the T-length history of previous observations and actions H_T, along with A_T, the action chosen at time T. Let $\mathbb{E}_{H_T, A_T \sim \Pi_T}[\mathcal{R}(H_T, A_T)]$ denote the expectation over the random variables $H_T = \langle G_0, A_1, .., A_{T-1}, G_T\rangle$ and A_T, where A_i are chosen according to $\Pi_T(H_i)\,\forall\, i \in [1, T]$, and G_i are drawn according to the distribution over uncertain edges of G_{i-1} that are revealed by A_i. The objective of DIME is to find an optimal T-step policy $\boldsymbol{\Pi_T^*} = \text{argmax}_{\Pi_T} \mathbb{E}_{H_T, A_T \sim \Pi_T}[\mathcal{R}(H_T, A_T)]$.*

6. DIME POMDP FORMULATION

DIME is modeled as a POMDP [17] (similar to [29]) because of two reasons. First, POMDPs are a good fit for DIME as (i) we conduct several interventions sequentially, similar to sequential POMDP actions; and (ii) we have *partial observability* (similar to POMDPs) due to uncertainties in network structure and influence status of nodes. Second, POMDP solvers have recently shown great promise in generating near-optimal policies efficiently [25]. We now explain how we map DIME onto a POMDP.

A POMDP is a tuple $\wp = \langle \mathbf{S}, \mathbf{A}, \mathbf{O}, \beta_0, \mathbf{T}, \boldsymbol{\Omega}, \mathbf{R}\rangle$, where $\mathbf{S}$, $\mathbf{A}$ and $\mathbf{O}$ are sets of possible world states, actions and observations respectively; β_0 is the initial belief state (distribution over states); $\mathbf{R}(\mathbf{s}, \mathbf{a}, \mathbf{s}')$ is the reward of taking action $\mathbf{a}$ in state $\mathbf{s}$ and reaching state $\mathbf{s}'$; $\mathbf{T}(\mathbf{s}'|\mathbf{s}, \mathbf{a})$ is the transition probability of reaching $\mathbf{s}'$ by taking action $\mathbf{a}$ in $\mathbf{s}$; $\boldsymbol{\Omega}(\mathbf{o}|\mathbf{a}, \mathbf{s}')$ is the observation probability of observing $\mathbf{o}$, by taking action $\mathbf{a}$ to reach $\mathbf{s}'$. We now explain how we map DIME onto a POMDP.

6.1 States

A POMDP state in our problem is a pair of binary tuples $s = \langle W, F\rangle$ where W and F are of lengths $|V|$ and $|E_U|$, respectively. Intuitively, W denotes the influence status of network nodes, where $W_i = 1$ denotes that node i is influenced and $W_i = 0$ otherwise. Moreover, F denotes the existence of uncertain edges, where $F_i = 1$ denotes that the i^{th} uncertain edge exists in reality, and $F_i = 0$ otherwise (assuming we order the nodes and uncertain edges). For example, in Figure 5, if *only* node A is influenced, and *only* uncertain edge (A, B) exists, then the POMDP state $s = \langle W, F\rangle$ is given by $W = \langle 1, 0, 0, 0, 0, 0\rangle$, because only node A is influenced (i.e, $W_1 = 1$) and all other nodes are un-influenced (i.e, $W_i = 0$); and $F = \langle 1, 0, 0, 0\rangle$ because out of the four uncertain edges in Figure 5, only (A, B) exists ($F_1 = 1$) and the other uncertain edges don't exist ($F_i = 0$). Thus, the set of all possible POMDP states are all possible combinations of the binary vectors W and F. We denote the set of all possible POMDP states by S.

6.2 Actions

Every choice of a subset of K nodes is a POMDP action. More formally, $A = \{a \subset V s.t. |a| = K\}$. For example, in Figure 5, one possible action is $\{A, B\}$ (when $K = 2$). We denote the set of all possible POMDP actions by A.

6.3 Observations

Upon taking a POMDP action, we "*observe*" the ground reality of the uncertain edges outgoing from the nodes chosen in that action. Consider $\Theta(a) = \{$e | e = (x,y) s.t. x $\in a\ \wedge$ e $\in E_u\}\ \forall a \in A$, which represents the (ordered) set of uncertain edges that are observed when we take action a. Then, our POMDP observation upon taking action a is defined as $o(a) = \{F_e | e \in \Theta(a)\}$, i.e., the F-values of the observed uncertain edges. For example, by taking action $\{B, C\}$ in Figure 5, the values of F_4 and F_5 (i.e., the F-values of uncertain edges in the 1-hop social circle of nodes B and C) would be observed. We denote the set of all possible POMDP observations by O.

6.4 Rewards

The reward $R(s, a, s')$ of taking action a in state s and reaching state s' is the number of newly influenced nodes in s'. More formally, $R(s, a, s') = (\|s'\| - \|s\|)$, where $\|s'\|$ is the number of influenced nodes in s'.

6.5 Initial Belief State

The initial belief state is a distribution β_0 over all states $s \in S$. The support of β_0 consists of all states $s = \langle W, F \rangle$ s.t. $W_i = 0 \,\forall\, i \in [1, |V|]$, i.e., all states in which all network nodes are un-influenced (as we assume that all nodes are un-influenced to begin with). Inside its support, each F_i is distributed independently according to $P(F_i = 1) = u(e)$. Recall that despite this assumption, there is uncertainty in the influence status of nodes in future time steps, because HEALER does not observe the nodes that have been influenced in between interventions. The only information HEALER has is that explicitly chosen nodes (i.e., our intervention participants in all past rounds) are influenced.

6.6 Transition And Observation Probabilities

Computation of exact transition probabilities $T(s'|s, a)$ requires considering all possible paths in a graph through which influence could spread, which is $\mathcal{O}(N!)$ (N is number of nodes in the network) in the worst case. Moreover, for large social networks, the size of the transition and observation probability matrix is prohibitively large (due to exponential sizes of state and action space).

Therefore, instead of storing huge transition/observation matrices in memory, we follow the paradigm of large-scale online POMDP solvers [25, 6, 5] by using a generative model $\Lambda(s, a) \sim (s', o, r)$ of the transition and observation probabilities. This generative model allows us to generate on-the-fly samples from the exact distributions $T(s'|s, a)$ and $\Omega(o|a, s')$ at very low computational costs. Given an initial state s and an action a to be taken, our generative model Λ simulates the random process of influence spread to generate a random new state s', an observation o and the obtained reward r. Simulation of the random process of influence spread is done by "*playing*" out propagation probabilities (i.e., flipping weighted coins with probability $p(e)$) according to our influence model to generate sample s'. The observation sample o is then determined from s' and a. Finally, the reward sample $r = (\|s'\| - \|s\|)$ (as defined above). This simple design of the generative model allows significant scale and speed up (as seen in previous work [25] and also in our experimental results section).

Next, we give a high-level overview of HEAL algorithm. For more detailed understanding, please refer to [29].

7. HEAL

HEAL solves the *original POMDP* using a novel *hierarchical ensembling heuristic*: it creates ensembles of imperfect (and smaller) POMDPs at *two* different layers, in a hierarchical manner (see Figure 6). HEAL's *top layer* creates an ensemble of smaller sized *intermediate POMDPs* by subdividing the original *uncertain network* into several smaller sized *partitioned networks* by using graph partitioning techniques [12]. Each of these partitioned networks is then mapped onto a POMDP, and these *intermediate POMDPs* form our *top layer* ensemble of POMDP solvers.

In the bottom layer, each *intermediate POMDP* is solved using TASP (**T**ree **A**ggregation for **S**equential **P**lanning), our novel POMDP planner, which subdivides the POMDP into another ensemble of smaller sized *sampled POMDPs*. Each member of this *bottom layer* ensemble is created by randomly sampling uncertain edges of the partitioned network to get a sampled network having no uncertain edges, and this sampled network is then mapped onto a *sampled POMDP*. Finally, the solutions of POMDPs in both the *bottom* and *top layer* ensembles are aggregated using novel techniques to get the solution for HEAL's original POMDP.

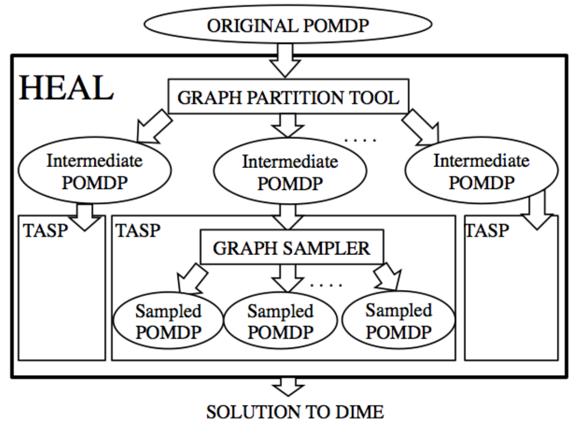


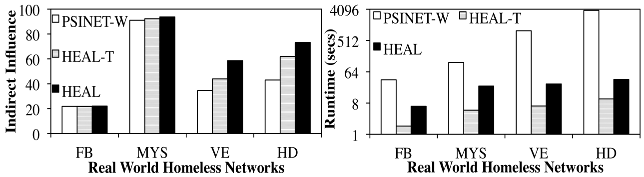
Figure 6: Hierarchical decomposition in HEAL

HEAL uses several novel heuristics. First, it uses a novel two-layered *hierarchical ensembling heuristic*. Second, it uses graph partitioning techniques to partition the uncertain network, which generates partitions that minimize the edges going across partitions (while ensuring that partitions have similar sizes). Since these partitions are "almost" disconnected, we solve each partition separately. Third, it solves the *intermediate POMDP* for each partition by creating smaller-sized *sampled POMDPs* (via sampling uncertain edges), each of which is solved using a novel tree search algorithm, which avoids the exponential branching factor seen in PSINET [30]. Fourth, it uses novel aggregation techniques to combine solutions to these smaller POMDPs rather than simple plurality voting techniques seen in previous ensemble techniques [30].

These heuristics enable scale up to real-world sizes (at the expense of sacrificing performance guarantees), as instead of solving one huge problem, we now solve several smaller problems. However, these heuristics perform very well in practice. Our simulations show that even on smaller settings, HEAL achieves a 100X speed up over PSINET, while providing a 70% improvement in solution quality; and on larger problems, *where PSINET is unable to run at all*, HEAL continues to provide high solution quality.

8. EXPERIMENTAL RESULTS

In this section, we analyze HEAL's performance on some settings. Both our experiments are run on a 2.33 GHz 12-core Intel machine having 48 GB of RAM. All experiments are averaged over 100 runs. We use a metric of "*Indirect Influence*" throughout this section, which is number of nodes "*indirectly*" influenced by intervention participants. For example, on a 30 node network, by selecting 2 nodes each for 10 interventions (horizon), 20 nodes (a lower bound for any strategy) are influenced with certainty. However, the total number of influenced nodes might be 26 (say) and thus, the *Indirect Influence* is $26 - 20 = 6$. In all experiments, the propagation and existence probability values on all network edges were uniformly set to 0.1 and 0.6, respectively. This was done based on findings in Kelly et. al.[8]. However, we relax these parameter



(a) Solution Quality (b) Runtime

Figure 7: Solution Quality and Runtime on Real World Networks

settings later in the section. All experiments are statistically significant under bootstrap-t ($\alpha = 0.05$). For a more comprehensive set of results, please refer to [29].

Baselines: We use two algorithms as baselines. We use PSINET-W as a benchmark as it is the most relevant previous algorithm, which was shown to outperform heuristics used in practice; however, we also need a point of comparison when PSINET-W does not scale. No previous algorithm in the influence maximization literature accounts for uncertain edges and uncertain network state in solving the problem of sequential selection of nodes; in-fact we show that even the standard Greedy algorithm [9, 7] has no approximation guarantees as our problem is not adaptive submodular. Thus, we modify Greedy by replacing our uncertain network with a certain network (in which each uncertain edge e is replaced with a certain edge e_0 having propagation probability $p(e_0) = p(e) \times u(e)$), and then run the Greedy algorithm on this *certain network*. We use the Greedy algorithm as a baseline as it is the best known algorithm known for influence maximization and has been analyzed in many previous papers [2, 1, 27, 9, 13, 7].

Datasets: We use *four real world social networks* of homeless youth, provided to us by our collaborators. All four networks are friendship based social networks of homeless youth living in different areas of a big city in USA (name withheld for anonymity). The first and second networks are of homeless youth living in two large areas (denoted by VE and HD to preserve anonymity), respectively. These two networks (each having ∼150-170 nodes, 400-450 edges) were created through surveys and interviews of homeless youth (conducted by our collaborators) living in these areas. The third and fourth networks are relatively small-sized online social networks of these youth created from their Facebook (34 nodes, 120 edges) and MySpace (107 nodes, 803 edges) contact lists, respectively. When HEALER is deployed, we anticipate even larger networks, (e.g., 250-300 nodes) than the ones we have in hand and we also show run-time results on artificial networks of these sizes.

Solution Quality/Runtime Comparison. We compare *Indirect Influence* and run-times of HEAL, HEAL-T and PSINET-W on all four real-world networks. We set $T = 5$ and $K = 2$ (since PSINET-W fails to scale up beyond $K = 2$ as shown later). Figure 7a shows the *Indirect Influence* of the different algorithms on the four networks. The X-axis shows the four networks and the Y-axis shows the *Indirect Influence* achieved by the different algorithms. This figure shows that (i) HEAL outperforms all other algorithms on every network; (ii) *it achieves ∼70% improvement over PSINET-W* in VE and HD networks; (iii) it achieves ∼25% improvement over HEAL-T. The difference between HEAL and other algorithms is not significant in the Facebook (FB) and MySpace (MYS) networks, as HEAL is already influencing almost all nodes in these two relatively small networks. Thus, in experiments to come, we focus more on the VE and HD networks.

Figure 7b shows the run-time of all algorithms on the four networks. The X-axis shows the four networks and the Y-axis (in log scale) shows the run-time (in seconds). This figure shows that (i) *HEAL achieves a 100X speed-up over PSINET-W*; (ii) PSINET-W's run-time increases exponentially with increasing network sizes; (iii) HEAL runs 3X slower than HEAL-T but achieves 25% more *Indirect Influence*. Hence, HEAL is our algorithm of choice that we plan to deploy in our pilot study. Next, we report on initial progress made in the pilot study.

9. PILOT STUDY WITH HOMELESS YOUTH

We now discuss ongoing efforts towards deploying HEALER in collaboration with Safe Place for Youth in a pilot study. This study will serve as a precursor to a much larger study where we plan to enroll 900 youth into our program. For our pilot study, we have begun generating the network using HEALER's Facebook application.

So far, we have enrolled 60 homeless youth into our pilot study. Over a period of two weeks, each youth that visited Safe Place for Youth was asked about the possibility of them enrolling into our study. Upon their consent, they were explained the goal and reason behind conducting this pilot study - to raise awareness about HIV in their social circles. Each youth was gifted a 20 US dollar gift card for enrolling into the study. They will also be given 25 and 30 US dollar gift cards for showing up after one and three months for follow-up interviews which will be used to assess influence spread. Finally, the youth were also given a three digit personal identification number (PID) using which they will be referred to in the pilot study (as part of an Institutional Review Board requirement to protect the anonymity of homeless youth at all times).

Using this PID, they logged into HEALER's Facebook application and then the social network was generated. Currently, the network is in the process of being refined with suggestions made by officials at Safe Place for Youth. Figure 8 shows a portion of the raw network generated using HEALER's Facebook application. Each node shows the PID of a homeless youth. Note that even though Facebook friendships are mostly bidirectional, we have replaced those bidirectional edges with two unidirectional edges (in order to account for the asymmetry of influence propagation in either direction of most friendships). After this network is refined with suggestions from officials at Safe Place for Youth, link prediction techniques will be used to infer missing or uncertain edges. Once this entire process is over, HEALER will be used to generate recommendations for homeless shelter officials.

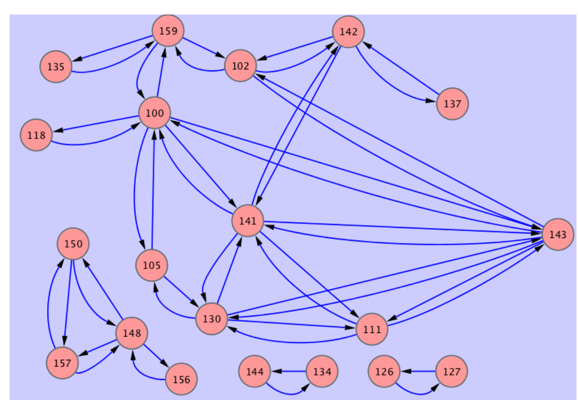


Figure 8: A portion of the raw network generated using HEALER's Facebook application

10. EXPLANATION SYSTEM

In our initial talks with homeless shelter officials, we observed that the shelter officials found HEAL's solutions to be very counter-intuitive. This was to be expected since people generally tend to perform bad in situations where multi-step expected utility calculations need to be done in order to find the best action (e.g., sequential planning problems). Since HEAL will be deployed among homeless youth whose welfare is the responsibility of the homeless shelter, we want the shelter officials to be comfortable with solutions provided by HEAL.

Our goal is to be able to justify the solutions of HEAL to the homeless shelter officials in an intuitive manner. This goal is slightly different than the goal of "*explaining*" HEAL's solution to the shelter official. Explaining HEAL's solution would entail telling the official exactly how HEAL calculates its solution, and then explain why HEAL chooses a particular choice of nodes in the network. That is, we would need to give a "*correct*" explanation to the official which would involve maximum expected utility calculations. On the other hand, we just want to justify the solutions of HEAL to the official in an intuitive manner. This means that we want to explain the solution of HEAL in a way that does not go against the official's intuition. In such a case, the official will be much more comfortable in adopting HEAL's solutions.

One possible way of designing this explanation system is to ensure that our system refrains from using MEU (maximum expected utility) calculations to justify HEALER's solutions. Instead, it could explain the solution in terms of concepts that the officials finds believable (or concepts that mirror the officials' intuition). For example, the officials might pick nodes which are centrally located and highly popular in the network. Now, degree centrality is not necessarily an optimal strategy, but if we can explain HEAL's solution in terms of "centrality and popularity" of nodes, then the official might be more willing to agree with the POMDP solution.

To that end, our first goal is to find out what kind of reasoning do officials (or humans in general) use to pick nodes in very simple graph settings. That will give us an understanding about what kinds of reasons are most likely to persuade humans and offials to adopt HEAL's solutions. Next, we describe the details of an Amazon Mechanical Turk (AMT) game that we have developed in order to find the biases/reasons that humans use to pick nodes.

10.1 Mechanical Turk Game

Our Amazon Mechanical Turk game collects data from human subjects which will help us understand the reasons using which people select nodes in networks. Our game is comprised of two different phases.

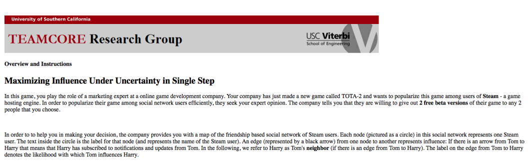


Figure 9: Instructions Page of AMT Game

10.1.1 *First Phase of Game*

In the first phase, our game collects data from human subjects by showing them pictures of different graphs and asking them to pick nodes in those graphs. Our game has four different variants, each of which is designed to gauge difficulties faced by humans in different settings. The four settings are as follows:

- **Short + Certain**: In this setting, subjects are asked to select two nodes for a single horizon on 8 different graphs with certain propagation of influence on all edges.

- **Short + Uncertain**: In this setting, subjects are asked to select two nodes for a single horizon on 8 different graphs with uncertain propagation of influence on all edges.

- **Long + Certain**: In this setting, subjects are asked to select two nodes for a two rounds on 8 different graphs with certain propagation of influence on all edges.

- **Long + Uncertain**: In this setting, subjects are asked to select two nodes for a two rounds on 8 different graphs with uncertain propagation of influence on all edges.

Data collected from these four variants will help us understand where do humans fail. Specifically, it will help us distinguish between whether humans fail at lookahead search (**Long + Certain** and **Long + Uncertain**) or at expected utility calculations (**Short + Uncertain**). At the beginning of the game, each human subject is randomly assigned to one out of the four possible game variants. He/she is shown the first set of graphs and his responses are recorded.

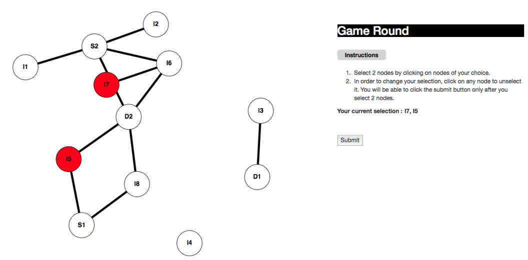


Figure 10: Short + Certain Variant

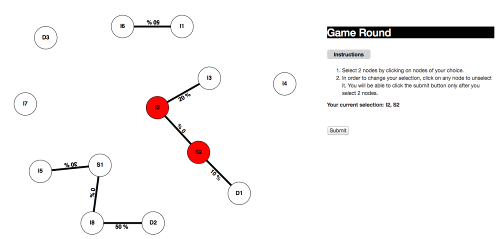


Figure 11: Short + Uncertain Variant

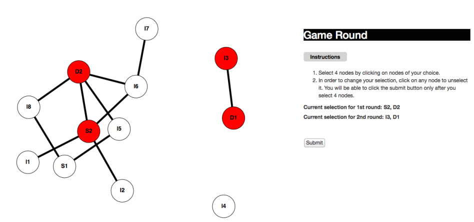


Figure 12: Long + Certain Variant

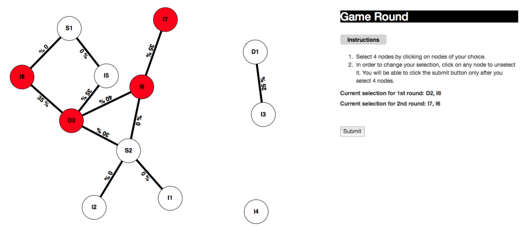


Figure 13: Long + Uncertain Variant

10.1.2 Second Phase of Game

The second phase of the game collects data on whether people find it easier to verify correct solutions as opposed to coming up with correct solutions. Recall that in the first phase of our game, each human subject was shown a set of eight different networks and they were asked to select a set of nodes for maximizing influence spread. In the second phase, for each of these eight networks, the human subject is shown four different solutions for that network. The four different solutions are as follows: (i) his own solution from Phase 1; (ii) HEALER's solution; (iii) solution based on Degree Centrality (i.e., nodes picked in order of decreasing degree centrality); (iv) solution based on Betweenness Centrality (i.e., nodes picked in order of decreasing betweenness centrality).

Data collected from the second phase will help us in finding out if people can verify the correct solutions (i.e., HEALER's solutions) even though it might be harder for them to come up with the correct solutions (due to the various difficulties that we test in the first phase of the game). For example, if we find out that people mostly figure out that HEALER's solutions are better than the other solutions, then the need for our POMDP explanation system would be negated. Moreover, if people do not select their own first phase solution in the second phase, that would point to the fact that peoples' biases towards their particular solution are not that strong.

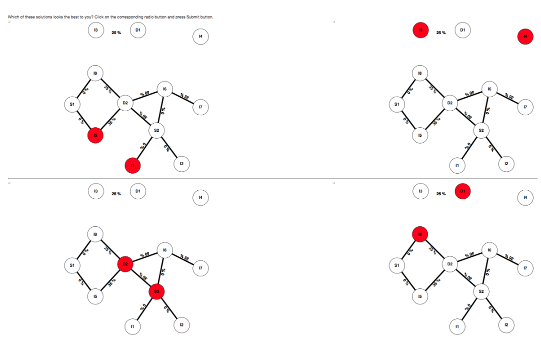


Figure 14: Second phase of AMT game

Compensation Scheme: Each human subject receives a base compensation of 50 cents and gets a bonus amount proportional to his/her performance on the task (the performance is judged by how close his solution is to the optimal solution). The bonus amount is capped off at one dollar, i.e., if the subject selects the optimal solution on each of the eight networks, he/she gets 1.5 dollars (one dollar bonus + 50 cent base compensation). Data collection using the game is currently underway[2]. In future work, we plan to utilize the collected data to guide the development of our POMDP persuasion system.

[2]The game can be played at http://cs-server.usc.edu:16292/

11. CONCLUSION

This paper looks at challenges faced during the ongoing deployment of HEALER, a POMDP based software agent that recommends sequential intervention plans for use by homeless shelters, who organize these interventions to raise awareness about HIV among homeless youth. HEALER's sequential plans (built using knowledge of social networks of homeless youth) choose intervention participants strategically to maximize influence spread, while reasoning about uncertainties in the network. In order to compute its plans, HEALER (i) casts this influence maximization problem as a POMDP and solves it using a novel planner which scales up to previously unsolvable real-world sizes; (ii) and constructs social networks of homeless youth at low cost, using a Facebook application. HEALER is currently being deployed in the real world in collaboration with a homeless shelter. Initial feedback from the shelter officials has been positive but they were surprised by the solutions generated by HEALER as these solutions are very counter-intuitive. Therefore, there is a need to justify HEALER's solutions in a way that mirrors the officials' intuition. In this paper, we report on progress made towards HEALER's deployment and detail first steps taken to tackle the issue of explaining HEALER's solutions. Specifically, we build a game on Amazon Mechanical Turk to collect data from human subjects in order to understand biases and reasons that humans use to select nodes in networks. This is the first step towards building our explanation system that will justify solutions of HEALER to homeless shelter officials in an intuitive manner.

12. ACKNOWLEDGEMENTS

This research was supported by MURI Grant W911NF-11-1-0332 and NIMH Grant number R01-MH093336.

REFERENCES

[1] C. Borgs, M. Brautbar, J. Chayes, and B. Lucier. Maximizing Social Influence in Nearly Optimal Time. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 946–957. SIAM, 2014.

[2] E. Cohen, D. Delling, T. Pajor, and R. F. Werneck. Sketch-based Influence Maximization and Computation: Scaling up with guarantees. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 629–638. ACM, 2014.

[3] J.-P. Cointet and C. Roth. How Realistic Should Knowledge Diffusion Models Be? *Journal of Artificial Societies and Social Simulation*, 10(3):5, 2007.

[4] N. H. Council. HIV/AIDS among Persons Experiencing Homelessness: Risk Factors, Predictors of Testing, and Promising Testing Strategies. `www.nhchc.org/wp-content/uploads/2011/09/InFocus_Dec2012.pdf`, Dec. 2012.

[5] J. S. Dibangoye, G. Shani, B. Chaib-Draa, and A.-I. Mouaddib. Topological Order Planner for POMDPs. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2009.

[6] A. Eck and L.-K. Soh. To Ask, Sense, or Share: Ad Hoc Information Gathering. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 367–376. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[7] D. Golovin and A. Krause. Adaptive Submodularity: Theory

and Applications in Active Learning and Stochastic Optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011.

[8] J. A. Kelly, D. A. Murphy, K. J. Sikkema, T. L. McAuliffe, R. A. Roffman, L. J. Solomon, R. A. Winett, and S. C. Kalichman. Randomised, Controlled, Community-Level HIV-Prevention Intervention for Sexual-Risk Behaviour among Homosexual men in US cities. *The Lancet*, 350(9090):1500, 1997.

[9] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the Spread of Influence through a Social Network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[10] O. Z. Khan, P. Poupart, and J. P. Black. Minimal Sufficient Explanations for Factored Markov Decision Processes. Citeseer, 2009.

[11] M. Kim and J. Leskovec. The Network Completion Problem: Inferring Missing Nodes and Edges in Networks. In *Proceedings of the SIAM Conference on Data Mining*. SIAM, 2011.

[12] D. LaSalle and G. Karypis. Multi-threaded Graph Partitioning. In *Parallel & Distributed Processing (IPDPS), 2013 IEEE 27th International Symposium on*, pages 225–236. IEEE, 2013.

[13] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance. Cost-effective Outbreak Detection in Networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 420–429. ACM, 2007.

[14] L. Marcolino, A. Lakshminarayanan, A. Yadav, and M. Tambe. Simultaneous Influencing and Mapping Social Networks. In *Proceedings of the Fiftienth International Conference on Autonomous Agents and Multiagent Systems (Short Paper) (AAMAS 2016)*, 2016.

[15] S. Paquet, L. Tobin, and B. Chaib-Draa. An Online POMDP Algorithm for Complex Multiagent Environments. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 970–977. ACM, 2005.

[16] P. Poupart, K.-E. Kim, and D. Kim. Closing the Gap: Improved Bounds on Optimal POMDP Solutions. In *International Conference on Automated Planning and Scheduling*, 2011.

[17] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2009.

[18] E. Rice. The Positive Role of Social Networks and Social Networking Technology in the Condom-using Behaviors of Homeless Young People. *Public health reports*, 125(4):588, 2010.

[19] E. Rice, A. Barman-Adhikari, N. G. Milburn, and W. Monro. Position-specific HIV risk in a Large Network of Homeless Youths. *American journal of public health*, 102(1):141–147, 2012.

[20] E. Rice, A. Fulginiti, H. Winetrobe, J. Montoya, A. Plant, and T. Kordic. Sexuality and Homelessness in Los Angeles public schools. *American Journal of Public Health*, 102, 2012.

[21] E. Rice, E. Tulbert, J. Cederbaum, A. B. Adhikari, and N. G. Milburn. 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, 2012.

[22] S. Ross, J. Pineau, S. Paquet, and B. Chaib-Draa. Online Planning Algorithms for POMDPs. *Journal of Artificial Intelligence Research*, pages 663–704, 2008.

[23] J. A. Schneider, A. N. Zhou, and E. O. Laumann. A new HIV Prevention Network Approach: Sociometric Peer Change Agent Selection. *Social Science & Medicine*, 125:192–202, 2015.

[24] B. Seegebarth, F. Müller, B. Schattenberg, and S. Biundo. Making hybrid plans more clear to human users-a formal approach for generating sound explanations. In *Twenty-Second International Conference on Automated Planning and Scheduling*, 2012.

[25] D. Silver and J. Veness. Monte-Carlo Planning in large POMDPs. In *Advances in Neural Information Processing Systems*, pages 2164–2172, 2010.

[26] M. T. Spaan and N. Vlassis. Perseus: Randomized Point-based Value Iteration for POMDPs. *Journal of artificial intelligence research*, pages 195–220, 2005.

[27] Y. Tang, X. Xiao, and Y. Shi. Influence maximization: Near-Optimal Time Complexity meets Practical Efficiency. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 75–86. ACM, 2014.

[28] T. W. Valente and P. Pumpuang. Identifying Opinion Leaders to Promote Behavior Change. *Health Education & Behavior*, 2007.

[29] A. Yadav, H. Chan, A. Jiang, H. Xu, E. Rice, and M. Tambe. Using Social Networks to Aid Homeless Shelters: Dynamic Influence Maximization under Uncertainty - An Extended Version. In *Proceedings of the Fiftienth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, 2016.

[30] A. Yadav, L. Marcolino, E. Rice, R. Petering, H. Winetrobe, H. Rhoades, M. Tambe, and H. Carmichael. Preventing HIV Spread in Homeless Populations Using PSINET. In *Proceedings of the Twenty-Seventh Conference on Innovative Applications of Artificial Intelligence (IAAI-15)*, 2015.

[31] Q. Yan, S. Guo, and D. Yang. Influence Maximizing and Local Influenced Community Detection based on Multiple Spread Model. In *Advanced Data Mining and Applications*, pages 82–95. Springer, 2011.

[32] S. D. Young and E. Rice. Online Social Networking Technologies, HIV knowledge, and Sexual Risk and Testing Behaviors among Homeless Youth. *AIDS and Behavior*, 15(2):253–260, 2011.