Optimizing HIV Interventions for Multiplex Social Networks Via Partition-Based Random Search

Qingpeng Zhang [Member, IEEE],
Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong SAR, China, and Shenzhen Research Institute of City University of Hong Kong, Shenzhen, Guangdong, China, e-mail: qingpeng.zhang@cityu.edu.hk.

Lu Zhong,
Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong SAR, China, e-mail: luzhong2-c@my.cityu.edu.hk.

Siyang Gao [Member, IEEE], and
Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong SAR, China, e-mail: siyangao@cityu.edu.hk.

Xiaoming Li
Arnold School of Public Health, University of South Carolina, Columbia, SC, e-mail: xiaoming@mailbox.sc.edu.

Abstract

There are multiple modes for HIV transmissions, each of which is usually associated with a certain key population (e.g. needle sharing among people who inject drugs). Recent field studies revealed the merging trend of multiple key populations, making HIV intervention difficult because of the existence of multiple. In this research, we aim to address this challenge by developing a multiplex social network framework, we propose a new random search method, named Partition-based Random Search with Network and Memory Prioritization (PRS-NMP), to identify the optimal subset of high-value individuals in the social network for interventions. Numerical experiments demonstrated that the proposed PRS-NMP based interventions could effectively reduce the scale of HIV transmissions. The performance of PRS-NMP based interventions is consistently better than benchmark Nested Partitions method and network-based metrics.

Keywords

Social networks; HIV transmissions; infectious disease; simulation optimization; partition-based random search

I. INTRODUCTION

HIV (Human immunodeficiency virus) is a major healthcare burden worldwide [1]. There are multiple modes for HIV transmissions, such as unprotected sex, sharing needles, mother-to-infant transmission, and transfusing blood, etc. [2]. Behavioral research has recognized that the interactions between multiple key populations could facilitate the transmission of HIV [3]–[6]. Recent field studies found that many female sex workers (FSWs) tended to

have noncommercial partners that are persons who inject drugs (PWID) [7]–[9]. This overlap results in an elevated risk of HIV outbreaks [10], [11]. It is critical to develop effective intervention strategies to reduce the risk of HIV transmissions via identifying the optimal subset of high-value individuals in the social network.

Social network has been widely adopted to characterize and model the epidemics of HIV [11]–[16] and other infectious diseases [13], [17], [18]. Behavioral research with real-world data found that an individual's social networking behaviors (e.g. exchanges of resources) underlied and sustained his or her HIV-risk behaviors [19]. Due to the limitations in performing large-scale experiments with human subjects, simulation has been a major tool to study HIV interventions on social networks. Real world social networks usually exhibit scalefree property [20], [21], causing difficulties in acquiring global immunity with randomly immunized individuals. However, their vulnerability to epidemic attacks can be sharply lowered by targeted immunization/intervention schemes [18], [22], [23].

Existing studies mainly represent the relationships between people with a simplex framework (e.g. unprotected sex) [24], [25]. However, there are usually multiple types of relationships through which infectious diseases transmit [2]. Field studies revealed that different transmission modes exhibit different transmission patterns of HIV, making the conventional simplex network framework inadequate to model the epidemics patterns accurately [7], [26], [27]. In order to model the multiple types of relationships through which HIV transmits among FSWs and PWID, Zhong *et al.* developed a multiplex network based model in a simulation study, and found that bridge individuals (who have connections in multiple key populations) played a significant role in HIV transmissions [11]. In addition, the authors simulated the interventions based on individuals' positions in the multiplex social network, and demonstrated that preventions on bridge individuals could help reduce both the scale and speed of HIV transmissions. However, this preliminary study only examined the intervention strategies that isolated individuals based on a single network property (degree, betweenness centrality, or cross-layer betweenness centrality). These network-based intervention methods are usually not optimal. Research on the optimization of HIV interventions for such multiplex social networks is still yet to come. In addition, selecting the optimal subset of individuals in network based models is usually NP-hard [28], [29], making the search for high-value individuals for targeted HIV intervention very challenging.

In this paper, we propose a new random search method, named Partition-based Random Search with Network and Memory Prioritization (PRS-NMP) to address this challenge. The problem is based on a multiplex social network framework, in which each node represents a person (either in FSWs community or PWID community), and each edge represents the relationship between the two corresponding nodes (either unprotected sex or needle sharing). the optimal PRS-NMP aims to identify subset of *n* individuals to be isolated from the rest of the multiplex social network, so that the scale of HIV transmission is minimized.

The contribution of this research is twofold. First, we develop a multiplex social network framework to model the HIV transmissions, and simulate the outcomes of different intervention strategies. This framework is able to capture the multimode transmissions of

infectious diseases. It can also be easily applied to modeling other scenarios like marketing, opinion dynamics, information diffusion, social recommendations, etc. [30]–[33]. Second, we propose a novel networkbased partition random search method that can identify the best subset of individuals to form the optimal intervention strategy. The newly proposed method takes advantage of the network structure and the memory of prior partitions to facilitate the search for optimal solution. It can be expanded to solve other optimization problems in a network context, such as influence maximization, social synchrony, control of complex networks, resource allocation, traffic control etc. [28], [34]–[40].

The remaining of this paper is organized as follows. Section II introduces the background and preliminaries. Section III describes the proposed PRS-NMP method. Section IV presents results of numerical experiments. We conclude the paper with discussions of limitations and future work in Section V.

II. BACKGROUND AND PRELIMINARIES

In this section, we first introduce the multiplex social network framework for modeling HIV transmissions across key populations. Then, the preliminaries of partition-based random search framework are presented.

A. Multiplex Social Network Framework for Modeling HIV Transmissions across FSWs and PWID Communities

We construct a multiplex social network framework to model the HIV transmissions across two key populations – FSWs and PWID. The two populations form two communities, which are connected through a small number of crosscommunity *bridge* edges, representing the non-commercial partnerships between FSWs and PWID. Figure 1 presents a hypothetical example of the multiplex social network framework. In this example, there is a bridge formed by one FSW node and one PWID node (the two adjacent nodes are *bridge nodes*). The bridge could be formed by unprotected sex relationship between the two nodes, or sharing needles while taking drugs (in this case, both). If the initial infected node is in the FSWs community, the HIV virus would be transmitted to PWID community through the bridge nodes, and vice versa.

The network is presented as $G = (V_1, V_2, E_1, E_2)$, where V_1 is the set of nodes in the FSWs community, and V_2 is the set of nodes in the PWID community. $V = V_1 \cup V_2$ is the set of all nodes. E_1 and E_2 represent the *unprotected sex relationships* e_{ij}^1 and *needle sharing relationships* e_{ij}^2 among nodes, respectively. $E = E_1 \cup E_2$ is the set of all edges. As both unprotected sex and sharing needles relationships are mutual, the network is undirected, $e_{ij} \equiv e_{ji}$. The set of *bridge* edges is represented by $B = \{e_{ij}^1 | v_i \in V_1, v_j \in V_2\} \cup \{e_{ij}^2 | v_i \in V_1, v_j \in V_2\}$. Bridge edges are incident to *bridge nodes* $C = \{v_i | e_{ij}^1 \in B\} \cup \{v_i | e_{ij}^2 \in B\}$. Both the FSWs community and PWID community were found to be scale-free networks, in which a *rich-gets-richer effect* exists [24], [41], [42]. Therefore, we adopted the classic preferential attachment model to generate the social networks of both communities [20], [21]. FSWs community consists of FSWs and their regular clients. PWID community consists of PWID

who share needles with others while using drugs. We assume that there is no homosexual relationship among them, thus the FSWs community a bipartite network. We modified the classic preferential attachment model to create a bipartite network with the scale-free property to represent the FSWs community.

In the numerical experiments, we constructed one multiplex social network as the platform for simulations, according to the descriptive data we obtained from interviewing local FSWs and PWID in a certain province in Southwest China [7]. The network has 500 nodes in each community with 5% nodes randomly selected as bridge nodes. Figure 2 shows the visualization of the network. As shown by the topological properties in Table I and the degree distribution in Figure 3, both communities follow a heavy-tail degree distribution (largely power-law) with relatively short distance between each other. It is worth noting that due to the bipartite property enforced in the modified preferential attachment model, the degree distribution of FSWs community does not follow a power-law perfectly when the degree is small. These properties are aligned with the empirical studies in the literature and our prior field studies [24], [26], [42].

B. Background of Partition-based Random Search

The partition-based random search (PRS) is a popular optimization scheme that iteratively partitions the solution space into a number of subregions, evaluates how promising each subregion is, and then moves to the most promising subregion to continue the partitioning operations. PRS has been recognized as an efficient heuristic scheme in both deterministic and stochastic optimization problems [43]–[48]. Below we summarize the basic framework of PRS:

- *Partitioning*. The solution space is denoted by Θ. In the initial step, the solution space is partitioned into τ subregions $\sigma_1, \sigma_2, \ldots, \sigma_\tau$. In the latter partitioning operations, the partitioning of the solution space is based on the quality of subregions obtained from past sampled solutions, so that computational effort is mainly spent in areas with high quality.

- *Random sampling*. Solution points are generated through sampling in each partitioned regions.

- *Evaluating the Promising Index* and *Backtracking*. N_i points are sampled from region σ_i. The promising index $I(\sigma_i)$ of region σ_i is evaluated based on sampled points.

In this step, one region shall be determined as the most promising region $\sigma_{\hat{j}}$ for the next iteration. Then, go back to the first step for further partitioning iterations.

There are many variants of PRS with different partitioning, sampling, promising index evaluation, and backtracking methods. In this research, we proposed a novel method based on the classic Nested Partitions method [45].

III. METHODOLOGY

We develop a novel partitioning scheme, named Partitionbased Random Search with Network and Memory Prioritization (PRS-NMP) for the search for high-value individuals in social networks. There are two main innovations: First, while partitioning the solution space, PRS-NMP prioritizes the nodes with higher value of a certain network property. In this study, the prioritization is on nodes with high degree and high betweenness centrality, because these two properties have been recognized to be strongly associated with the connectivity of nodes [11], [21], [49]. Therefore, individuals located at the core of the community are critical for the robustness and resilience of the whole network [49]–[51]. Second, PRSNMP records the information of partitions in each iteration, and backtracks to a smaller promising region (where the best solution exists) for next iteration. This memory-based backtracking method takes advantage of previous partitioning information to improve the efficiency of the original PRS methods. In the context of HIV intervention, the problem is to identify the best subset of nodes to be taken away (isolated), so that the total number of infected individuals can be minimized in the social network.

A. Problem Formulation

Given a network G described in Section II.A, we consider the following optimization problem:

$$\min_{\theta^K \in \Theta^K} f\left(\theta^K\right) \quad (1)$$

where Θ^K is the solution set for K isolated nodes. $\# V$ is the number of nodes in the network, so $\# \Theta^K = \binom{\# V}{K}$. θ^K is the solution for a fixed value of K. The number of nodes in each solution $\theta^K = \{s_k\} \subseteq V$ is K, where $k = 1, 2,...,K$, and $1 \leq K \leq \# V$. $f: \Theta^K \to \mathbb{R}$ is the objective function to be optimized. Though this is a simple search problem, the huge number of alternatives makes the enumeration approach infeasible. Most problems can only guarantee global optimum through checking all the alternatives. The search for an optimal subset of nodes in complex networks is usually NP-hard [28], [29].

We assume that as long as a node can be reached by an infected node through a connected path in the social network, the node is a *risk node* with a chance of getting infected. The nodes that cannot be reached by infected nodes are *riskfree* nodes. Here, we randomly select initially infected nodes. Therefore, the problem is to identify the best subset of K nodes to be isolated, so that the expected number of *risk nodes* is the lowest, or, the expected number of *risk-free* nodes is the highest. It is essentially a cut problem. For each solution θ^K, the original network G is cut into a number of connected components $G_{\theta^K}(c)$, where $1 \leq c \leq \# V$. We aim to minimize the expected number of nodes that are connected to the initially infected nodes within a connected component. The objective function is then expressed as:

$$f\left(\theta^K\right) = \sum_{i=1}^{c} p(i) \,\#\, G_{\theta^K}(i), \quad (2)$$

$$p(i) = 1 - \frac{\binom{\#\, G - K - \#\, G_{\theta^K}(i)}{H}}{\binom{\#\, G - K}{H}}, \quad (3)$$

where H is the number of initially infected nodes, which are randomly selected from all nodes with equal likelihood. $\#G_{\theta}K(i)$ is the number of nodes in the ith connected component generated by cutting the original network G with θ^K. $p(i)$ represents the probability of nodes in $G_{\theta}K(i)$ being selected as the initially infected nodes. In this paper, we mainly investigate the scenarios in which $H = 1, 2$, thus

$$p(i)\big|_{H=1} = \frac{\#\, G_{\theta^K}(i)}{\#\, G - K}, \quad (4)$$

$$p(i)\big|_{H=2} = \frac{(\#\, G - K)^2 - \left[\#\, G - K - \#\, G_{\theta^K}(i)\right]^2 - \#\, G_{\theta^K}(i)}{(\#\, G - K)^2 - \#\, G + K} \quad (5)$$

$$= \frac{-\#\, G_{\theta^K}(i)^2 + 2\, \#\, G\, \#\, G_{\theta^K}(i) - 2K\, \#\, G_{\theta^K}(i) - \#\, G_{\theta^K}(i)}{\#\, G^2 - 2\, \#\, GK + K^2 - \#\, G + K}.$$

In the context of multiplex network, the problem formulation is more complicated, but follows the same framework. In particular, we investigate the HIV interventions in three scenarios:

- *S-S scenario*. There are two initially infected nodes, both in the FSWs community. In this case, the expression of $p(i)$ is:

$$p(i)|_{H=2} = \begin{cases} 0 & G_{\theta^K}(i) \cap V'_1 = \varnothing \\ 1 - \dfrac{\binom{\# V'_1 - \# G_{\theta^K}{}^{(i)}}{2}}{\binom{\# V'_1}{2}} & G_{\theta^K}(i) \cap V'_1 = G_{\theta^K}(i) \\ 1 - \dfrac{\binom{\# V'_1 - \# G^1_{\theta^K}{}^{(i)}}{2}}{\binom{\# V'_1}{2}} & G^1_{\theta^K}(i) \subseteq G_{\theta^K}(i) \end{cases} \qquad (6)$$

where $V'_1 = V_1 \backslash V_1 \cap \theta^K$ represents the nodes in the FSWs community after intervention, $G^1_{\theta^K}(i) = G_{\theta^K}(i) \cap V'_1$ represents the FSWs community nodes that are in $G_{\theta^K}(i)$.

- *D-D scenario*. There are two initially infected nodes, both in the PWID community. In this case, the expression of $p(i)$ is:

$$p(i)|_{H=2} = \begin{cases} 0 & G_{\theta^K}(i) \cap V'_2 = \varnothing \\ 1 - \dfrac{\binom{\# V'_2 - \# G_{\theta^K}{}^{(i)}}{2}}{\binom{\# V'_2}{2}} & G_{\theta^K}(i) \cap V'_2 = G_{\theta^K}(i) \\ 1 - \dfrac{\binom{\# V'_2 - \# G^2_{\theta^K}{}^{(i)}}{2}}{\binom{\# V'_2}{2}} & G^2_{\theta^K}(i) \subseteq G_{\theta^K}(i) \end{cases} \qquad (7)$$

where $V'_2 = V_2 \backslash (V_2 \cap \theta^K)$ represents the nodes in the PWID community after intervention, $G^2_{\theta^K}(i) = G_{\theta^K}(i) \cap V'_2$ represents the PWID community nodes that are in $G_{\theta^K}(i)$.

- *S-D scenario*. There are two initially infected nodes, one in the PWID community and the other one in the PWID community. In this case, the expression of $p(i)$ is:

$$p(i)\big|_{H=2} = 1 - \frac{\left(\# V_1' - \# G^1_{\theta^K}(i)\right)\left(\# V_2' - \# G^2_{\theta^K}(i)\right)}{\# V_1' \, \# V_2'}, \quad (8)$$

B. PRS-NMP Algorithm

The proposed PRS-NMP is based on the classic Nested Partitions method for global optimization developed by Shi and Ólafsson [45]. Since PRS-NMP prioritizes topologically important nodes and takes advantage of previously identified promising region, it is expected to identify the optimal solution more efficiently. First, we set K (the number of isolated nodes) as a fixed value. Then, there are four main steps:

Step 1. Prioritization.—Before any iteration, we prioritize the nodes according to their values of a certain network property (e.g. degree, betweenness centrality). The *rank* of a node $rank(V_i) \in \mathbb{N}^+$ is in descendant order by the value of the network property. The node with the largest value has the highest rank. For example, if the degree of v_1 is 20, and the degree v_2 is 2, then the rank of v_1 is *higher* than the rank of v_2, denoted as $rank(v_1) <$ $rank(v_2)$. If two or more nodes have the same value, their ranks are determined arbitrarily.

Step 2. Partitioning.—For each iteration r, the current most promising region is denoted as $\sigma^{K,r,*} \in \Theta^K$. The surrounding region $\Theta^K \backslash \sigma^{K,r,*}$ is denoted as $\sigma_3^{K,r}$. Thus, $\sigma^{K,1,*} = \Theta^K$, $\sigma_3^{K,1} = \varnothing$. We define the rank of a region $rank(\sigma)$ as the highest *rank* of all the nodes in the region. Thus, $rank(\sigma^{K,1,*}) = rank(\Theta^K) = 1$. Then, we partition $\sigma^{K,r,*}$ into 2 subregions $\sigma_1^{K,r}$ and $\sigma_2^{K,r} \cdot \sigma_1^{K,r}$ represents the subregion of equal rank, $rank(\sigma_1^{K,r}) = rank(\sigma^{K,r,*})$. $\sigma_2^{K,r}$ represents the subregion of a rank lower than the rank of the current most promising region.

Step 3. Random Sampling.—We randomly sample N_j solutions from region $\sigma_j^{K,r}$, and then calculate the outcome of the objective function for each solution

$$f\left(\theta_j^{K,r,1}\right), f\left(\theta_j^{K,r,2}\right), ..., f\left(\theta_j^{K,r,N_j}\right), \quad j = 1,2,3.$$

In our study, we set $N_1 = N_2 = ... = N_j = 100$. With each region, all the solutions have equal positive probability to be sampled.

Step 4. Evaluating the Promising Index and Backtracking.—We estimate the performance of each sampled solution as follows:

$$\hat{I}\left(\sigma_j^{K,r}\right) = \min_{i \in \left\{1,2,...,N_j\right\}} f\left(\theta_j^{K,r,i}\right) \quad j = 1,2,3. \quad (9)$$

Then, we need to determine the most promising region $\sigma_{\hat{j}}$, where

$$\hat{j} \in \min_{j} \hat{I}\left(\sigma_j^{K,r}\right) \qquad j = 1, 2, 3. \quad (10)$$

If more than one regions are equally promising, we arbitrarily choose one of them as the most promising region. If the newly obtained most promising region is a subregion of $\sigma_1^{K,r}$ or $\sigma_2^{K,r}$, we further partition the corresponding region. In the PRS-NMP method, we record the information of partitions in each iteration. One solution only belongs to one region. If a region is partitioned into several subregions, the old region is discarded. Each solution in the discarded region belongs to one of the newly partitioned subregions. The union of all the current subregions is the solution space Θ. Therefore, if the newly obtained best solution is in $\sigma_3^{K,r}$ (not in $\sigma^{K,r,*}$), we determine which specific subregion (partitioned in previous iterations) the solution belongs to, and set that subregion as the new most promising region. We can then go back to Step 2 to continue the partitioning and sampling.

Stopping Criteria.: The algorithm stops if the current nonpartitionable solution is the best so far, and the algorithm does not backtrack to other regions for the most recent 100 iterations consecutively. Then, the output is the current nonpartitionable solution.

For the HIV intervention problem, we developed four variants of the PRS-NMP method, and three benchmark methods:

- *Degree*. We isolate the nodes with largest values of degree. Degree is the number of edges incident to the node. It has been recognized as a critical metric to evaluate the importance of a node in diffusing viruses. Nodes with larger degree tend to be the hubs of the community.

- *Betweenness*. We isolate the nodes with largest values of betweenness centrality. Betweenness centrality is the proportion of shortest path traversing through a node. It has been recognized as a critical metric to evaluate the importance of a node in diffusing viruses. Nodes with larger betweenness centrality tend to be those connecting different clusters in the network.

- *Nested Partitions*. This is the classic method on which the PRS-NMP method is based. We adopted algorithm AP 1 in [45].

- *PRS-NMP (D)*. We prioritize nodes according to their degrees in *Step 2*. We do not record the information of previously partitioned subregions in *Step 4*.

- *PRS-NMP (B)*. We prioritize nodes according to their betweenness centrality in *Step 2*. We do not record the information of previously partitioned subregions in *Step 4*.

- *PRS-NMP (DM)*. We prioritize nodes according to their degrees in *Step 2*. We record the information of previously partitioned subregions in *Step 4*.

- • *PRS-NMP (BM).* We prioritize nodes according to their betweenness centrality in *Step 2.* We record the information of previously partitioned subregions in *Step 4.*

IV. RESULTS

Based on the multiplex social network constructed in Section II.A, we consider three scenarios introduced in Section III – S-D, S-S, and D-D. For each scenario, we generated 1000 problems through randomly selecting initially infected nodes for 1000 times. In total, we have 3000 problems to solve. Using the four variants of PRS-NMP methods and three benchmark methods, we performed a set of numerical experiments with varying intervention budgets (ratio of nodes to be isolated in the network). The experiments were run in four parallel threads on a Dell T430 Tower Server with an Intel E5–2640v4 2.4GHz CPU and 80GB RAM.

Figure 4 presents the results of the experiments. In general, the more budget (higher ratio of isolated nodes) we had, the lower the ratio of *risk nodes* was. PRS-NMP based methods consistently performed better than degree/betweenness-based methods and the original Nested Partitions method. Note that the problems were generated randomly, thus we did not have ground truth of the global optimum. However, these results clearly demonstrated the effectiveness of the proposed PRS-NMP method in identifying an optimized set of highvalue individuals for HIV interventions. Specifically, the best performing method varied in different scenarios.

For S-D scenario, the resulted ratio of risk nodes for PRSNMP (DM) was the lowest (highest effectiveness). The other PRS-NMP based methods also performed well, except that PRS-NMP (B) generated less optimized results than those of network property based methods when the budget was20% or more. We also found that PRS-NMP (DM) and PRS-NMP (BM) performed better than PRS-NMP (D) and PRSNMP (B), indicating the huge benefits brought by the memory mechanism of the proposed methods. In addition, degree-based methods were more effective, indicating the more important role played by centrally connected nodes.

For S-S and D-D scenarios, there were critical segregation points. Because the initially infected nodes were in one community, achieving these segregation points indicated that the HIV viruses were limited in one community, instead of transmitting to both communities. The results showed that, other than the classic Nested Partitions, both the PRS-NMP based methods and degree/betweenness-based methods could achieve the segregation points. Interestingly, PRS-NMP (DM) segregated the network with the lowest budget level (12.5%) in the S-S scenario; whereas PRS-NMP (B) achieved so in the D-D scenario with the lowest budget level (10%) in the D-D scenario. The different performances are due to the behavioral difference between FSWs and PWID. Although with the same size and density, the FSWs community is a bipartite network (as we do not consider homosexual relationships) and PWID community is a simple scale-free network. The bipartite nature of the FSWs community makes the nodes with higher degrees even more important in serving as the hub of disease transmissions and, thus, leads to the better performance of degree-

based PRS-NMP method in the SS scenario. On the other hand, the existence of triangular relationships in PWID community weakens the importance of densely-connected nodes (high degree) and, thus, lead to the better performance of betweenness-based PRS-NMP method in the D-D scenario.

Comparing the results for the three scenarios, we found that the optimal solutions for D-D problems led to smaller ratios of risk nodes (higher effectiveness), and the segregation point could be reached with a lower budget level. This indicated that the HIV transmission in the FSWs community was relatively more difficult to constrain – the FSWs community was more vulnerable as compared with PWID community. Given that the network size of the FSWs community was the same as PWID community, this result echoed previous research on the controllability of bipartite network – it was found that bipartite network is easier to *control* (more vulnerable) [52]. Here the term *control* refers to effort to bring all nodes to the same state of *driver nodes* [34]. In this case, the role of initially infected nodes is similar as driver nodes. The aim of the optimization problem is to make the network less controllable (less vulnerable) through removing a subset of nodes. Because S-S problems set both initially infected nodes in the FSWs community, a bipartite network, it was expected that the effectiveness of intervention was lower than other scenarios.

The original Nested Partitions method did not converge to optimum for most cases. Note that the original Nested Partitions method is a less greedy method, thus it has a good chance to find the global optimum given sufficient iterations. However, the problems in this study are highly complicated, making the convergence process slow. Comparing the CPU time consumed by the random search methods (Table II), the original Nested Partitions method took much longer time (doubled, even tripled) to search for solutions, but still generated less optimized results. The PRS-NMP based methods have similar CPU time in all cases.

Figure 5 presents the average degree, betweenness centrality, and portion of bridge nodes of the solutions identified by different methods for each scenario. Solutions generated by the original Nested Partitions method did not contain sufficient high-degree and high-betweenness centrality nodes, indicating that it was difficult to capture the network structure with a simple random sampling approach. On the other hand, solutions generated by degree/betweenness-based methods comprise these nodes with highest degree/betweenness values. The proposed PRS-NMP methods identified better solutions that were not solely consisted of nodes with high degree/betweenness values. This verifies that pure networkbased interventions are usually not optimal. The proportion of bridge nodes varied in the solutions for different scenarios, and was smaller than what we expected intuitively. It turns out that we do not have to isolate all bridge nodes to segregate the network; instead, the isolation of part of the bridge nodes can effectively cut the bridges connecting the FSWs and PWID communities.

V. CONCLUSIONS

The merge between multiple key populations makes HIV intervention much more difficult. Modeling the relationships between these at-risk individuals as a multiplex social network is

of potential to help decision makers characterize the transmission patterns that cannot be captured by simplex social networks, and develop effective HIV intervention strategies.

This paper presents, to the best of our knowledge, the first optimization approach to identifying the optimal subset of individuals for intervention from multiplex social networks. Based on the classic Nested Partitions method, we develop a novel PRS-NMP method that prioritizes nodes with certain network properties and keeps the record of all partitions. The effectiveness and efficiency of this novel heuristic optimization method were demonstrated by a set of numerical experiments with three scenarios.

The newly proposed PRS-NMP method can be applied to other stochastic optimization problems, particularly those concerning the structure of a complex network. The method itself also has a lot of room for improvements. In particular, the partitioning, sampling schemes, and backtracking schemes can be modified according to the network structure. The current prioritization scheme is based on a deterministic rank of nodes. It can be extended to a stochastic prioritization scheme in which the rank of a node leads to a corresponding probability that determines the priority of the node. In addition, the current method only uses one network property for node prioritization. It would be interesting to explore if combining different properties could lead to more efficient algorithms.

The outcomes of this research have the potential to be translated into real value in public health practice. PRS-NMP can be easily implemented as a decision support systems to identify high-value individuals in key populations for timely interventions. It will help social workers optimize the use of limited intervention budget (e.g. HIV testing equipment, free condoms, education activities, etc.) to achieve the best intervention outcome.

In practice, front-line social workers and HIV researchers can use the PRS-NMP method to identify high-risk FSWs/PWID, and then inform them the risk of HIV infection caused by different modes. There are multiple ways to reduce their risk behaviors, including assisting FSWs/PWID to do HIV testing, providing free condoms and syringes/needles, and educating them the available healthcare services. Given the important role of identified high-risk FSWs/PWID, they can also help disseminate HIV prevention and intervention information to other at-risk people, like other FSWs and their clients. In addition to HIV interventions, the PRS-NMP method can be applied to other infectious diseases that have multiple modes of transmissions.

This research has several limitations. First, the current problem is to reduce the ratio of risk nodes. In reality, a risk node that is many hops away from infected nodes is actually with a very low risk of getting infected, even it can be reached through a connected path in the social network. Second, we did not consider the speed of HIV transmissions. The current research considers the optimal cut of networks through isolating (removing) nodes. This approach cannot capture detailed transmission patterns over time. In the future work, we plan to adopt epidemiological models (like the Susceptible Infected model) to simulate detailed transmission processes, in which each edge is associated with a transmission rate between the two corresponding nodes. We will then modify the experiment to examine the speed and scale of HIV transmissions in a more realistic setting, so that we can generate

actionable insights for public health practice. The adoption of epidemiological models can also reveal the behavioral differences between the two communities with a higher resolution. Last, we will examine the theoretical underpinnings of the proposed PRSNMP method, and perform experiments with larger scale and higher dimensionality.

Acknowledgments

Biography
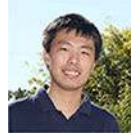
Qingpeng Zhang (M'10) received the B.S. degree in automation from the Huazhong University of Science and Technology, Wuhan, China, in 2009, and the Ph.D. degree in systems and industrial engineering from The University of Arizona, Tucson, AZ, in 2012. He was a Post-Doctoral Research Associate with The Tetherless World Constellation, Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY. He is currently an Assistant Professor with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong SAR, China. He is an Associate Editor of the IEEE Transactions on Computational Social Science and IEEE Transactions on Intelligent Transportation Systems. His research interests include social computing, complex networks, data mining, and semantic web.

Lu Zhong received the Bachelor of Science in Software Engineering from Southwest University, Chongqing, China, in 2015. She is currently pursuing the Ph.D. degree with in the Department of Systems Engineering and Engineering Management at City University of Hong Kong, Hong Kong SAR, China. Her research interests include healthcare analytics and multilayer network.

Siyang Gao (M'14) received the B.S. degree in mathematics from Peking University, Beijing, China, in 2009 and the Ph.D. degree in industrial engineering from University of Wisconsin-Madison, Madison, WI, in 2014. He is an Assistant Professor with the

Department of Systems Engineering and Engineering Management, City University of Hong Kong, China. His work has appeared in the IEEE Transactions on Automatic Control, Automatica, and the IEEE Transactions on Automation Science and Engineering. Dr. Gao is a member of INFORMS. His research is devoted to simulation optimization, large-scale optimization and radiation treatment.

Xiaoming Li is Professor and Endowed Chair of clinical translational research and the Director for the South Carolina SmartState Center for Healthcare Quality at University of South Carolina (USC) Arnold School of Public Health. As a psychologist with a very strong background in research methodology, intervention program development and cultural adaptation, clinical trial design, biostatistics, and psychometrics, he has been working with an interdisciplinary team of pediatricians, psychologists, and other health care professionals for over 24 years in the areas of health disparities and health promotion research in both domestic and international settings. He has been continuously funded by the World AIDS Foundation and NIH since 2000 to conduct research among children and adolescents, rural migrants, men who have sex with men, female sex workers, children and family affected by HIV/AIDS, and people living with HIV/AIDS in China. He was the founding editor for Health Psychology and Behavioral Medicine (from 2011–2016). His areas of research include HIV behavioral prevention, mental health, resilience-based psychosocial intervention, HIV disclosure, stigma reduction, and HIV treatment and care.

REFERENCES

[1]. WHO, UNICEF, and UNAIDS, "Global HIV/AIDS response- epidemic update and health sector progress towards universal access-progress report 2011," Tech. Rep, 2011 [Online]. Available: http://www.who.int/about/licensing/

[2]. Keeling RP, "HIV disease: current concepts," J. Couns. Dev, vol. 71, no. 3, pp. 261–274, 1993 [Online]. Available: http://doi.wiley.com/10.1002/j.1556-6676.1993.tb02212.x

[3]. Mathers BM, Degenhardt L, Ali H, Wiessing L, Hickman M, Mattick RP, Myers B, Ambekar A, and Strathdee SA, "HIV prevention, treatment, and care services for people who inject drugs: a systematic review of global, regional, and national coverage," Lancet, vol. 375, no. 9719, pp. 1014–1028, 2010. [PubMed: 20189638]

[4]. Merson MH, O'Malley J, Serwadda D, and Apisuk C, "The history and challenge of HIV prevention," Lancet, vol. 372, no. 9637, pp. 475–488, 2008. [PubMed: 18687461]

[5]. Strathdee SA, Philbin MM, Semple SJ, Pu M, Orozovich P, Martinez G, Lozada R, Fraga M, de la Torre A, Staines H, MagisRodríguez C, and Patterson TL, "Correlates of injection drug use among female sex workers in two Mexico-U.S. border cities," Drug Alcohol Depend, vol. 92, pp. 132–140, 2008. [PubMed: 17714888]

[6]. Strathdee SA, Lozada R, Martinez G, Vera A, Rusch M, Nguyen L, Pollini RA, Uribe-Salas F, Beletsky L, and Patterson TL, "Social and structural factors associated with HIV infection among female sex workers who inject drugs in the Mexico-US border region," PLoS One, vol. 6, no. 4, 2011.

[7]. Hong Y and Li X, "Behavioral studies of female sex workers in China: A literature review and recommendation for future research," AIDS Behav, vol. 12, no. 4, pp. 623–636, 2008. [PubMed: 17694431]

[8]. Saidel TJ, Des Jarlais D, Peerapatanapokin W, Dorabjee J, Singh S, and Brown T, "Potential impact of HIV among IDUs on heterosexual transmission in Asian settings: Scenarios from the Asian Epidemic Model," Int. J. Drug Policy, vol. 14, no. 1, pp. 63–74, 2003.

[9]. Ditmore MH, "When sex work and drug use overlap: considerations for advocacy and practice," Harm Reduction International, London, Tech. Rep., 2013.

[10]. Morris M and Kretzschmar M, "Concurrent partnerships and the spread of HIV," Aids, vol. 11, no. 5, pp. 641–648, 1997 [Online]. Available: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-199705000-00012 [PubMed: 9108946]

[11]. Zhong L, Zhang Q, and Li X, "Modeling the intervention of HIV transmission across intertwined key populations," Sci. Rep, vol. 8, 2018.

[12]. Hyman JM and Stanley E, "Using mathematical models to understand the AIDS epidemic," Math. Biosci, vol. 90, no. 1–2, pp. 415–473, 1988 [Online]. Available: http://www.sciencedirect.com/science/article/pii/0025556488900788

[13]. Meyers LA, Pourbohloul B, Newman ME, Skowronski DM, and Brunham RC, "Network theory and SARS: predicting outbreak diversity," J. Theor. Biol, vol. 232, no. 1, pp. 71–81, 2005. [PubMed: 15498594]

[14]. Needle RH, Coyle SL, Genser SG, and Trotter RT, "Social networks, drug abuse, and HIV transmission," NIDA Res. Monogr, vol. 151, 1995.

[15]. Newman ME, "Spread of epidemic disease on networks," Phys. Rev. E, vol. 66, 2002.

[16]. Toit D and Craig IK, "Selective pinning control of the average disease transmissibility in an hiv contact network," Phys. Rev. E, vol. 92, 2015 [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.92.012810

[17]. Parra-Rojas C, House T, and McKane AJ, "Stochastic epidemic dynamics on extremely heterogeneous networks," Phys. Rev. E, vol. 94, 2016 [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.94.062408

[18]. Wang JB, Wang L, and Li X, "Identifying spatial invasion of pandemics on metapopulation networks via anatomizing arrival history," IEEE Transactions on Cybernetics, vol. 46, no. 12, pp. 2782–2795, 2016. [PubMed: 26571544]

[19]. Johnson BT, Redding CA, Diclemente RJ, Mustanski BS, Dodge B, Sheeran P, Warren MR, Zimmerman RS, Fisher WA, Conner MT, Carey MP, Fisher JD, Stall RD, and Fishbein M, "A network-individual-resource model for HIV prevention," AIDS Behav, vol. 14, no. SUPPL. 2, pp. 204–221, 2010. [PubMed: 20862606]

[20]. Barabási AL and Albert R, "Emergence of scaling in random networks," Science, vol. 286, pp. 509–512, 1999. [PubMed: 10521342]

[21]. Newman M, Barabási AL, Watts DJ, and Boguñá M, The Structure and dynamics of networks, 2011.

[22]. Pastor-Satorras R and Vespignani A, "Immunization of complex networks," Phys. Rev. E, vol. 65, no. 3, 2002.

[23]. Salathe M and Jones JH, "Dynamics and control of diseases in networks with community structure," Plos Comput. Biol, vol. 6, no. 4, 2010.

[24]. Liljeros F, Edling CR, Amaral LAN, Stanley HE, and Aberg Y, "The web of human sexual contacts," Nature, vol. 441, pp. 907–908, 2001 [Online]. Available: 10.1038/35082140

[25]. Pastor-Satorras R and Vespignani A, "Epidemic dynamics and endemic states in complex networks," Phys. Rev. E, vol. 63, 2001 [Online]. Available: 10.1103/PhysRevE.63.066117

[26]. Yang H, Li X, Stanton B, Liu H, Liu H, Wang N, Fang X, Lin D, and Chen X, "Heterosexual transmission of HIV in China: a systematic review of behavioral studies in the past two decades," Sex. Transm. Dis, vol. 32, no. 5, pp. 270–280, 2005 [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1791011&tool=pmcentrez&rendertype=abstract [PubMed: 15849527]

[27]. De Domenico M, Granell C, Porter MA, and Arenas A, "The physics of spreading processes in multilayer networks," Nat. Phys, vol. 12, no. 10, pp. 901–906, 2016 [Online]. Available: 10.1038/nphys3865

[28]. Kempe D, Kleinberg J, and Tardos É, "Maximizing the spread of influence through a social network," Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, vol. 11, no. 4, pp. 105–147, 2015 [Online]. Available: http://portal.acm.org/citation.cfm?doid=956750.956769

[29]. Liu W, Deng ZH, Cao L, Xu X, Liu H, and Gong X, "Mining top k spread sources for a specific topic and a given node," IEEE Transactions on Cybernetics, vol. 45, no. 11, pp. 2472–2483, 2015. [PubMed: 25494522]

[30]. Sahneh FD, Scoglio C, and Van Mieghem P, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," IEEE/ACM Trans. Netw, vol. 21, no. 5, pp. 1609–1620, 2013.

[31]. Battiston F, Nicosia V, and Latora V, "The new challenges of multiplex networks: Measures and models," Eur. Phys. J. Spec. Top, vol. 226, no. 3, pp. 401–416, 2017.

[32]. Mao M, Lu J, Zhang G, and Zhang J, "Multirelational social recommendations via multigraph ranking," IEEE Transactions on Cybernetics, vol. 47, no. 12, pp. 4049–4061, 2017. [PubMed: 28113690]

[33]. Nakatsuji M, Zhang Q, Lu X, Makni B, and Hendler JA, "Semantic social network analysis by cross-domain tensor factorization," IEEE Transactions on Computational Social Systems, vol. 4, no. 4, pp. 207–217, 2017.

[34]. Liu YY, Slotine JJ, and Barabási AL, "Controllability of complex networks," Nature, vol. 473, no. 7346, pp. 167–173, 2011. [PubMed: 21562557]

[35]. Park M, Lee SH, Kwon OM, and Seuret A, "Closeness-centralitybased synchronization criteria for complex dynamical networks with interval time-varying coupling delays," IEEE Transactions on Cybernetics, vol. 99, pp. 1–11, 2018.

[36]. Xuan Q, Zhang ZY, Fu C, Hu HX, and Filkov V, "Social synchrony on complex networks," IEEE Transactions on Cybernetics, vol. 48, no. 5, pp. 1420–1431, 2018. [PubMed: 28500015]

[37]. Kuhnle A, Alim MA, Li X, Zhang H, and Thai MT, "Multiplex influence maximization in online social networks with heterogeneous diffusion models," IEEE Transactions on Computational Social Systems, vol. 5, no. 2, pp. 418–429, 2018.

[38]. Li Y and Morse AS, "The power allocation game on a network: A paradox," IEEE/CAA Journal of Automatica Sinica, vol. 5, no. 4, pp. 771–776, 2018.

[39]. Li L, Lv Y, and Wang F, "Traffic signal timing via deep reinforcement learning," IEEE/CAA Journal of Automatica Sinica, vol. 3, no. 3, pp. 247–254, 2016.

[40]. Wang F, Zheng N, Cao D, Martinez CM, Li L, and Liu T, "Parallel driving in cpss: A unified approach for transport automation and vehicle intelligence," IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 4, pp. 577–587, 2017.

[41]. Brewer DD, Potterat JJ, Garrett SB, Muth SQ, Roberts JM, Kasprzyk D, Montano DE, and Darrow WW, "Prostitution and the sex discrepancy in reported number of sexual partners," Proc. Natl. Acad. Sci. U. S. A, vol. 97, no. 22, pp. 12385–12388, 2000. [PubMed: 11027304]

[42]. Hsieh C, Kovarik J, and Logan T, "How central are clients in sexual networks created by commercial sex?" Sci. Rep, vol. 4, no. 7540, 2014 [Online]. Available: http://www.nature.com/doifinder/10.1038/srep07540

[43]. Horst R, "General class of branch-and-bound methods in global optimization with some new approaches for concave minimization," J. Optim. Theory Appl, vol. 51, no. 2, pp. 271–291, 1986.

[44]. Tang ZB, "Adaptive partitioned random search to global optimization," IEEE Trans. Automat. Contr, vol. 39, no. 11, pp. 2235–2244, 1994.

[45]. Shi L and Olafsson S, "Nested partitions method for global optimization," Oper. Reserach, vol. 48, no. 3, pp. 390–407, 2000.

[46]. Pichitlamken J and Nelson BL, "A combined procedure for optimization via simulation," ACM Trans. Model. Comput. Simul, vol. 13, no. 2, pp. 155–179, 2003.

[47]. Chen W, Gao S, Chen CH, and Shi L, "An optimal sample allocation strategy for partition-based random search," IEEE Trans. Autom. Sci. Eng, vol. 11, no. 1, pp. 177–186, 2014.

[48]. Gao S and Chen W, "A partition-based random search for stochastic constrained optimization via simulation," IEEE Trans. Automat. Contr, vol. 62, no. 2, pp. 740–752, 2017.

[49]. Vespignani A, "Complex networks: The fragility of interdependency," Nature, vol. 464, no. 7291, pp. 984–985, 2010. [PubMed: 20393545]

[50]. Dong G, Gao J, Du R, Tian L, Stanley HE, and Havlin S, "Robustness of network of networks under targeted attack," Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys, vol. 87, no. 5, 2013.

[51]. Gao J, Barzel B, and Barabási AL, "Universal resilience patterns in complex networks," Nature, vol. 530, no. 7590, pp. 307–312, 2016 [Online]. Available: 10.1038/nature16948 [PubMed: 26887493]

[52]. Nacher JC and Akutsu T, "Structural controllability of unidirectional bipartite networks," Sci. Rep, vol. 3, 2013.
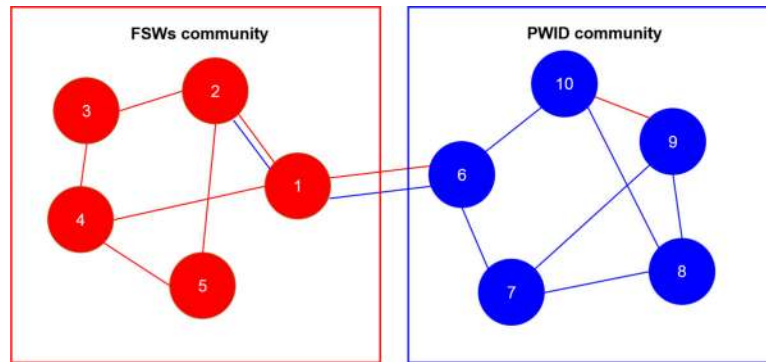
Fig. 1.
Illustration of the multiplex social network representing the relationships among individuals in two key populations. Red and blue nodes represent individuals in the FSWs community and PWID community, respectively. Red and blue edges represent the existence of unprotected sex and needle sharing behaviors, respectively. Node 1 and node 6 form noncommercial partnership, resulting in an elevated risk of HIV outbreaks across the two key populations.
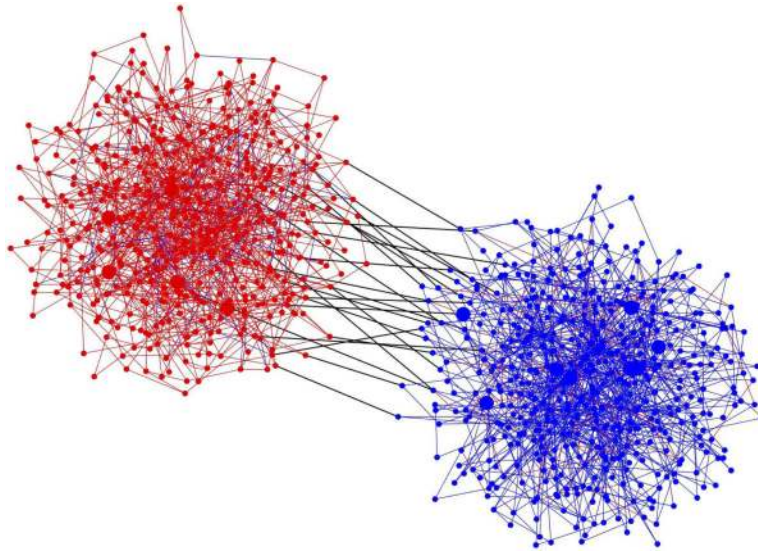
Fig. 2.
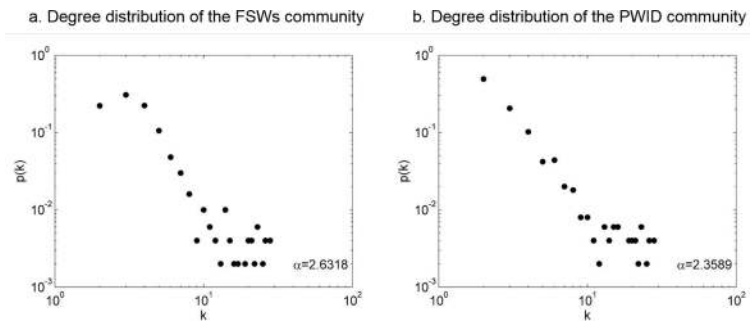Visualization of the multiplex social network consisted of FSWs (red) and PWID (blue).

Fig. 3.
Degree distributions of FSWs and PWID communities in the multiplex social network, and the fitted exponent in power-law distribution.
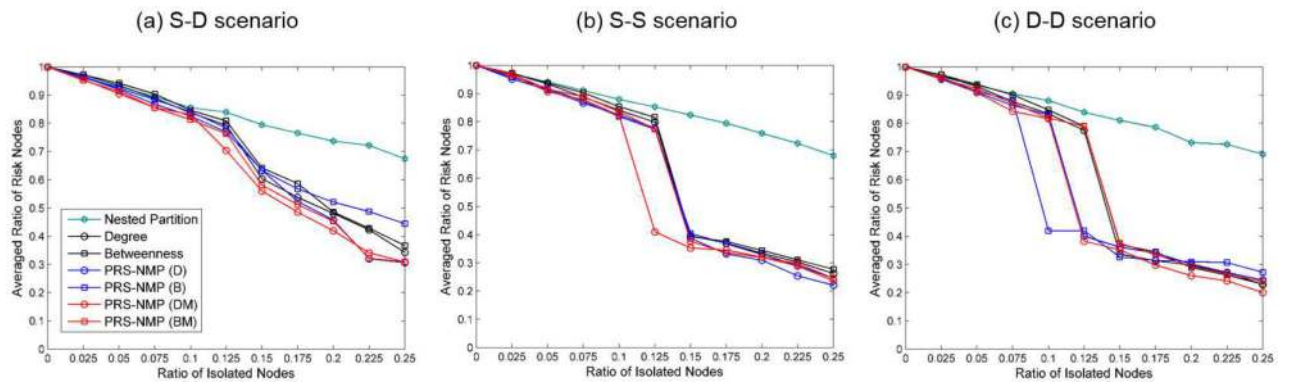
Fig. 4.
The ratio of risk nodes with varying intervention budgets and different intervention methods.
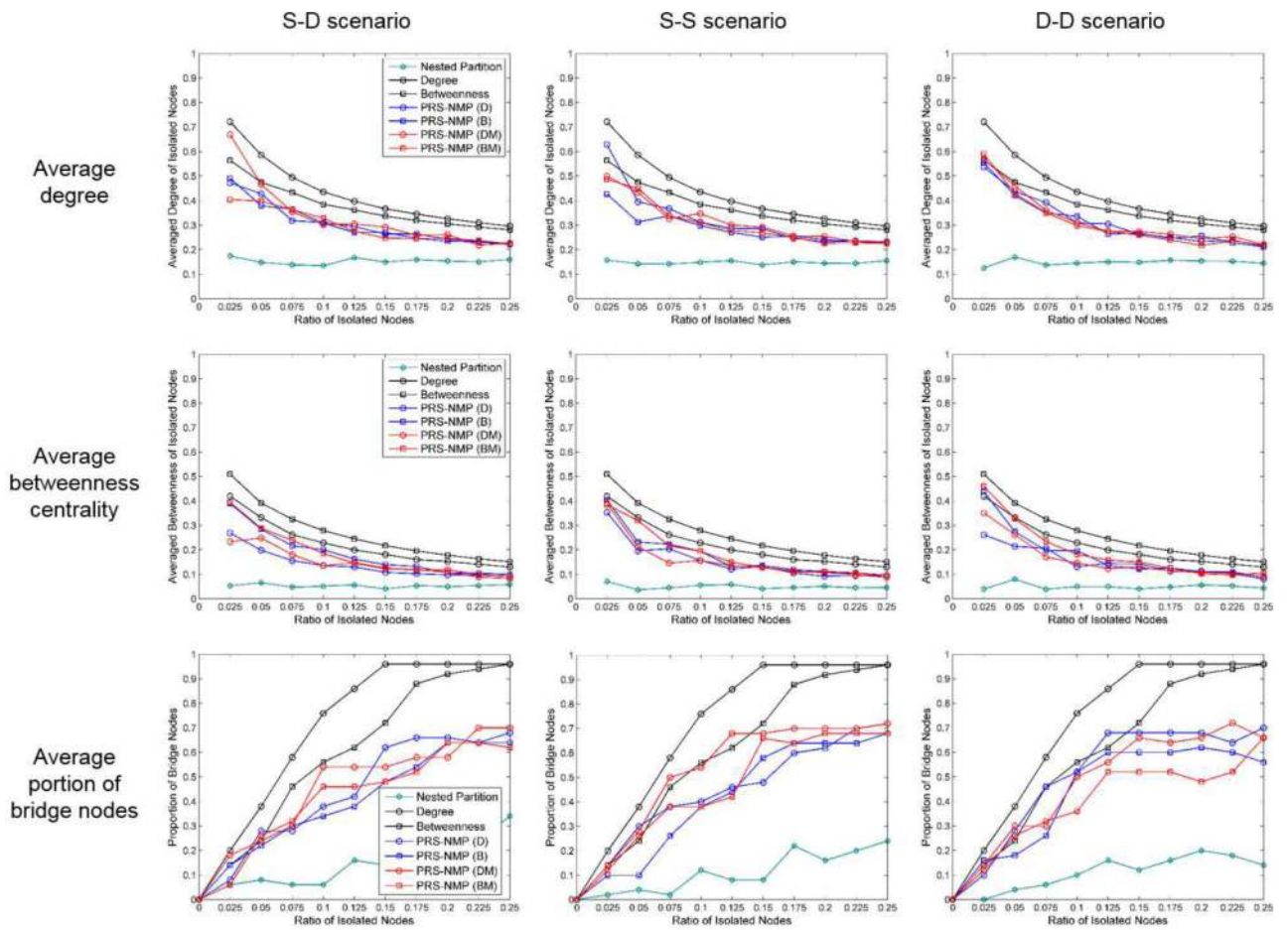
Fig. 5.
The average degree, betweenness centrality, and portion of bridge nodes of the solutions identified by different methods for each scenario.

Table I

TOPOLOGICAL PROPERTIES OF THE MULTIPLEX SOCIAL NETWORK.

Property	Value
Number of nodes in the FSWs community $|V_1|$	500
Number of nodes in the PWID community $|V_2|$	500
Number of edges in the FSWs community $|E_1|$	1000
Number of edges in the PWID community $|E_2|$	1000
Fraction of bridge nodes $|C|$	5%
Average degree $<k>$	4.50
Average shortest path length $<l>$	5.19
Average clustering coefficient $<c>$	0.02
Exponent in the power-law degree distribution of the FSWs community a_1	2.63
Exponent in the power-law degree distribution of the PWID community a_1	2.36
Diameter D	10

Table II

AVERAGE CPU TIME CONSUMED BY EACH METHOD FOR SOLVING A PROBLEM OF EACH SCENARIO.

Scenario	Avg. CPU time (s) Nested Partitions	Avg. CPU time (s) PRS-NMP (D)	Avg. CPU time (s) PRS-NMP (B)	Avg. CPU time (s) PRS-NMP (DM)	Avg. CPU time (s) PRS-NMP (BM)
S-D	8333	4301	4350	4575	4564
S-S	15922	4481	4710	4362	4590
D-D	8922	4200	4162	4493	4396