

Social Relation Based Long-term Vaccine Distribution Planning to Suppress Pandemic ^{*}

Donghyun Kim¹, Hao Guo¹, Yuchao Li², Wei Wang^{2**},
Sung-Sik Kwon¹, Alade O. Tokuta¹

¹ Dept. of Math. and Physics, North Carolina Central Univ., Durham, NC, USA.

{donghyun.kim, skwon, atokuta}@nccu.edu, hguo@eagles.nccu.edu

² School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China.

liyu.chao@126.com, wang.weiw@163.com

Abstract. This paper introduces a new optimization problem which aims to develop a distribution plan of vaccines which will be supplied over time such that an epidemic can be best suppressed until a complete cure for it is invented. We first exploit the concept of temporal graph to capture the projected images of the evolving social relations over time and formally define the social-relation-based vaccine distribution planning problem (SVDP²) on the temporal graph. Then, we introduce a graph induction technique to merge the subgraphs in the temporal graph into a single directed acyclic graph. Next, we introduce a max-flow algorithm based technique to evaluate the quality of any feasible solution of the problem. Most importantly, we introduce a polynomial time enumeration technique which will be used along with the evaluation technique to produce a **best possible** solution within polynomial time.

1 Introduction

In 2014, the world have witnessed the unprecedented spread of critical disease called Ebola, which is transmittable from an infected person, who has spent a certain incubation period after the initial infection, to another healthy one via direct contact to bodily fluid from the infected. After the seriousness of the

^{*} This work was supported in part by US National Science Foundation (NSF) CREST No. HRD-1345219. This research was jointly supported by National Natural Science Foundation of China under grants 11471005.

^{**} Corresponding Author

disease was recognized, many efforts were initiated to expedite the development of vaccines to stop the further spread of the disease as well as of a complete cure for it. During last one year, a number of vaccines were tested and several approaches to cure infected ones were tried, which saved several lives. However, the disease is still spreading and many people are dying while the researches for the vaccines and cures are ongoing. Unfortunately, there is no guarantee that Ebola is the last pandemic on this scale. As a result, the proper preparedness against such calamity is of great urgency to save lives, possibly in the near future.

It is not difficult to imagine that even a vaccine is once invented against a new pandemic, its near-term availability would be highly restricted. As a result, the development of proper distribution strategy of vaccines over healthy individuals is as important as inventing vaccines and cures of unknown pandemic to reduce the number of the victims of the critical epidemic once happens. Recently, Zhang and Prakash [1] used the information from social relationship to address the issue of selecting those to be vaccinated when the number of currently available vaccines is limited. In this approach, two adjacent nodes with high probability implies there is a great chance of infection from one to the other. This social relation based approach could be promising as the disease usually is transmitted one to another following their (physical) social interaction. However, we found that there is generally a lack of efforts to utilize this approach.

In order to fill this deficiency, this paper investigates the best way to distribute available vaccines which will be supplied over time by exploiting the projected social relations among the members of a society with the objective of minimizing the number of infected people until a complete cure is invented. The list of the contributions of this paper is as follows.

- (a) We introduce a new optimization problem, namely social-relation-based vaccine distribution planning problem (SVDP²), which aims to study the best strategy to distribute regular vaccine supplies over time with the objective of minimizing the number of infected until a complete cure is developed.
- (b) We use the concept of temporal graph [10] to capture the projected images of the evolving social relations over time. Then, we introduce a new strategy to reduce the graphs in the temporal graph into a single directed acyclic graph (DAG). Finally, we redefine the proposed optimization problem on this new DAG.

- (c) We introduce a new maximum-flow algorithm based strategy to evaluate the performance of any feasible solution of SVDP^2 .
- (d) We propose a polynomial time exact algorithm for SVDP^2 by exploiting our evaluation strategy.

The rest of this paper is organized as follows. Related work is discussed in Section 2. The formal definition of SVDP^2 is in Section 3. Our main contribution, the polynomial time exact algorithm for SVDP^2 is in Section 4. Finally, we conclude this paper in Section 5.

2 Related Work

We realize that there are several problems which are previously well-studied in the literature. Therefore, we need to explain how our problem is fundamentally different from them. Largely speaking, there are three group of problems related to ours.

Fire-fighter Problem and Its Variations [4–6]. There are two variations of the the fire-fighter problem [5]. In the context of our problem of interest, the first version, namely MAXSAVE , aims to find a valid vaccine strategy over time to maximize the number of uninfected after a given period. The second version, MINBUDGET , attempts to find a valid vaccine strategy to save the members in a given node subset with a given graph such that the budget for the vaccines (the number of nodes removed) used this purpose is minimized. At a glance, MAXSAVE is similar to our problem of interest, SVDP^2 . However, SVDP^2 is more challenging as MAXSAVE uses a static topology graph while SVDP^2 considers a social relation graph which varies over time.

Graph-cut Problems [2, 3]. The main objective of graph-cut problem is that given a graph, to identify a subset of nodes such that after the nodes in the subset are removed from the graph, the resulting graph consists of two connected components in a way that a certain objective function is maximized. One example of such objective is that each of the components should have one designated node s (and t), respectively and the size of the component including t becomes maximized. The main challenges to use a solution for the graph-cut problems for SVDP^2 are that (a) the former one assumes we have enough vaccines to contain the epidemic, which is not necessarily true in SVDP^2 , and (b) the former one also assumes a static graph, which is not necessarily true in our case.

Data-Aware Vaccination Problem [1]. Recently, Zhang and Prakash have investigated the data-aware vaccination problem, the problem of how to best distribute currently available k vaccines over healthy individuals so that the expected number of victims can be minimized with the knowledge of the infection probability from one to another under the assumption that infection of a patient to another happens only one time. In their work, the knowledge of the social network graph which represents the relationship between the people is used to evaluate the likelihood of the disease transmission. Then, a greedy strategy is used to find the best k healthy nodes in the graph such that the average number of patients are minimized. This work is very remote from our work as (a) there is no concept of time-dimension in their work, e.g. an infected individual may infect its neighbor only one time with a probability and the vaccines are only provided at the beginning, and (b) the social network graph is fixed.

Based on our survey, we can conclude that there is no existing work which is directly used to solve SVDP². In the following section, we provide the formal definition of SVDP².

3 Problem Definition

This paper uses a temporal graph [10] $\mathcal{G} = \{G_0 = (V_0, E_0), G_1 = (V_1, E_1), \dots, G_T = (V_T, E_T)\}$, where $G_t \in \mathcal{G}$ captures the social relation among the members of society at the t -th unit moment from the initial moment (0-th moment) to the final moment (T -th moment). After the final moment, it is highly anticipated that a complete cure of the disease will be developed. The time gap between two consecutive moments could range from an hour to weeks. For instance, in case of Ebola, usual incubation time is 2 weeks, and this can be used as a reasonable gap. From the initial relationship among the members in the society, and corresponding graph G_0 , the temporal graph \mathcal{G} can be generated by an existing strategy such as [8]. We assume there is a threshold to determine if there should be an edge between a pair of nodes at a moment, which implies that two members at the moment are close enough to infect each other in case that one of them is infected with very high probability. Note that the accuracy of this approach is out of the scope of this paper, and we simply assume that the algorithm used for this purpose is highly precise.

Now, due to the gravity, we list our main assumptions in more detail and corresponding justifications if necessary.

- (a) The temporal graph $\mathcal{G} = \{G_0 = (V_0, E_0), G_1 = (V_1, E_1), \dots, G_T = (V_T, E_T)\}$ representing the social relationship of the members of the society at each moment is known in advance, and is precise. After T unit moments later, a complete cure of the disease will be developed. Any G_i and G_{i+1} may differ in node set or edge set as the relationship can be highly dynamic.
- (b) The initial set of infected people I_0 in G_0 is known in advance. I_i will be used to represent the set of infected people in G_i .
- (c) After each unit moment from G_i to G_{i+1} , the neighbors of I_i in G_i will be infected in G_{i+1} . We argue that our approach considers the worst-case (in which the infection ratio from two people is 100% if they are related) and thus would be more rigorous to deal with a critical disease like Ebola rather than the probabilistic approach considered by Zhang and Prakash [1].
- (d) The initial vaccine supply $\mathcal{Q} = \{Q_0 = (p_0 = 0, q_0), Q_1 = (p_1, q_1), \dots, Q_l = (p_l, q_l)\}$ are known in advance, where $Q_i = (p_i, q_i) \in \mathcal{Q}$ is the information of i th vaccine supply and p_i is the arrival moment of q_i vaccines.
- (e) Shortly after time T , the complete cure for the disease will be developed.

Now, we provide the formal definition of our problem of interest.

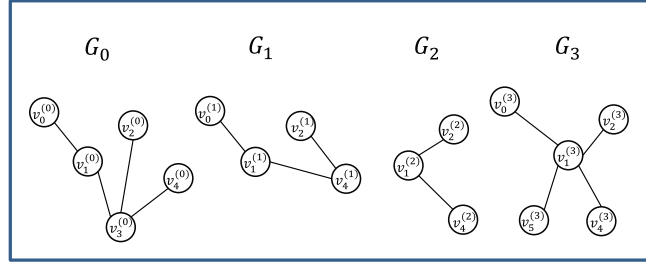
Definition 1 (Social-relation-based Vaccine Distribution Planning Problem (SVDP²)). *Given $\mathcal{G}, \mathcal{Q}, I_0$, and T , the goal of SVDP² is to find the best vaccine distribution schedule of the incoming vaccines under the infection model such that the total number of infected people after T unit moments is minimized.*

4 Main Contributions

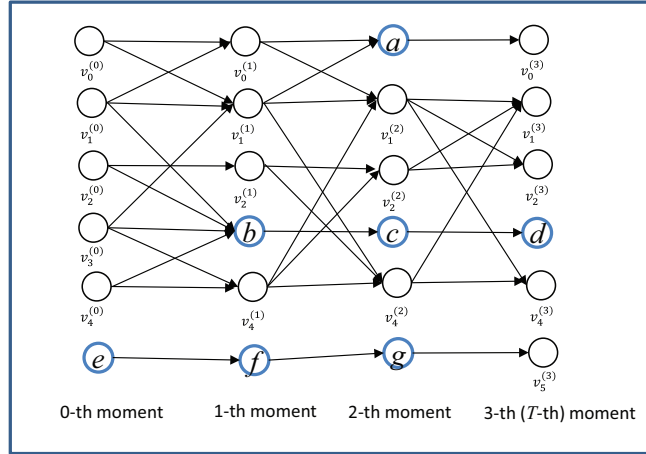
4.1 Consolidating \mathcal{G} to Integrated Graph \hat{G}

Apparently, \mathcal{G} is difficult to deal with as G_i and G_{i+1} in \mathcal{G} may differ in node sets and edge sets for any i . To overcome the difficulty, we introduce a graph consolidation technique to merge the graphs in \mathcal{G} to a new graph $\hat{G} = (\hat{V}, \hat{E})$, and redefine SVDP² using \hat{G} . This consists of the following steps.

- (a) **Node set construction:** Set $\hat{V} \leftarrow \bigcup_{0 \leq i \leq T} V(G_i)$, where $V(G_i)$ is the set of nodes in G_i . Each node $v_j^{(i)}$ represents the status of node v_j at the i -th



(a)



(b)

Fig. 1. (Figure (a) illustrates the time temporal graph \mathcal{G} which consists of a series of graphs G_1, G_2, G_3, G_4 representing the social relation at each moment. Figure (b) shows the integrated graph \hat{G} which is induced graph \mathcal{G} . In this graph, node a, b, c, d, e, f, g are fake nodes and does not exist. This means that e cannot be infected at the beginning.

moment. In case that there exists a node $v_j^{(i)} \in G_i$ for some i , but $v_j^{(k)} \notin G_k$ for some k , then add a virtual node $w_j^{(k)}$ to \hat{V} , e.g. nodes $\{a, b, c, d, e, f, g\}$ in Fig. 1(b).

- (b) **Edge set construction:** First, add a directed edge from $v_j^{(i)} \in \hat{V}$ (or alternatively $w_j^{(i)}$) to $v_j^{(i+1)} \in \hat{V}$ (or alternatively $w_j^{(i+1)}$) for each i and j pair: this means an infected node j at i -th moment will stay infected in $i + 1$ -th moment (even though the node is outside the area abstracted by the social network). Second, for each $v_j^{(i)} \in G_i$ and its neighbor $v_k^{(i)} \in G_i$, add a direct edge from $v_j^{(i)}$ to $v_k^{(i+1)}$ (or its virtual node $w_k^{(i+1)}$) to \hat{E} : this means that a

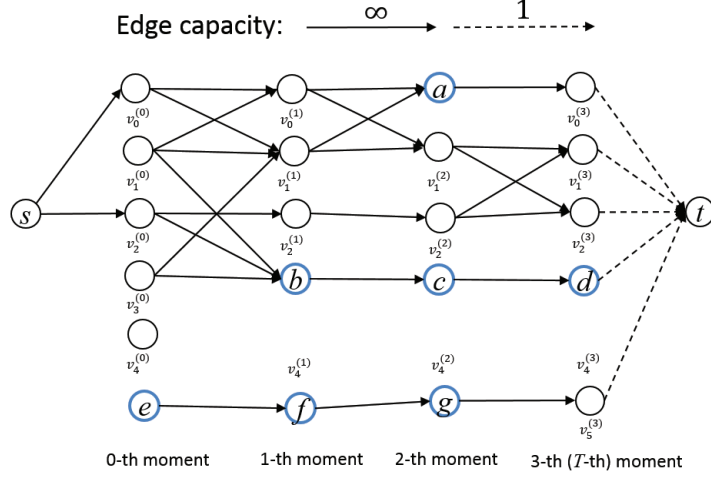


Fig. 2. In this graph, v_0 and v_2 are initially infected, and an available vaccine is given to v_4 at the initial moment. Then the maximum flow from s to t is equivalent to the number of infected nodes after the T -th moments, which is 4 in this example.

node neighboring to an infected node at i -th moment will keep infected in $i + 1$ -th moment.

The original infection rule can be applied in a way that when $v_j^{(i)}$ is infected, then its neighbors will be infected, and so on. Due to this, the resulting graph \hat{G} has the following two interesting property.

- (a) Once a node $v_j^{(i)}$ is infected, then all $v_j^{(k)}$ s such that $k > i$ will be infected.
- (b) Once a node $v_j^{(i)}$ is vaccinated, then all $v_j^{(k)}$ s such that $k > i$ will be vaccinated.

This means that once we decide a vaccine to $v_j^{(i)}$, then all of the nodes $v_j^{(k)}$ such that $k \geq i$ can be removed from \hat{G} . In fact, this is a unique property which distinguishes our problem with the rest of the existing related problems.

4.2 Evaluation of Feasible Solution

In this section, we introduce a max-flow algorithm based strategy to evaluate a feasible solution of SVDP². In detail, given an integrated graph $\hat{G} = (\hat{V}, \hat{E})$, we first add two nodes s, t to \hat{V} . Then, add an edge from s to the nodes in \hat{V} which

are initially infected. Then, set the edge capacity of all nodes in the current \hat{G} to be ∞ . Next, for all nodes at the T -th moment, we add an edge from each of them to t with an edge capacity 1 (see Fig. 3). Suppose the resulting graph is \hat{G}' . Then, we prove the following theorem.

Theorem 1. *The maximum $s-t$ flow in \hat{G}' after removing a subset of nodes S which received vaccines, i.e. if a node $v_k^{(i)} \in S$ receives a vaccine, then all nodes $v_k^{(j)}$ with $j \geq i$ and their corresponding edges are removed from \hat{G}' , is equivalent to the number of infected people after the T -th moment.*

Proof. Let I_T be the subset of nodes got infected in the final time $t = T$. Let f be a maximum flow with value $|f|$ for the constructed network in Fig. 2. We claim that for the maximum flow f , the flow must be one on the directed edge (v_i^T, t) for any $v_i^T \in I_T$. Otherwise, notice that for each $v_i^T \in I_T$, there is a directed path from s to v_i^T (since v_i^T got infected at moment T). We can take the directed path $(s \rightarrow v_i^T \rightarrow t)$ as a augmenting path, and increase the flow on every edge of the path by one (note the capacity of each edge from s to v_i^T is infinity, so we can increase the flow as we wish). Then the new flow value would increase by one; contradiction to the maximality of f . Therefore, we have $|f| = |I_T|$.

4.3 Polynomial Time Exact Algorithm based on Enumeration

In this section, we discuss how the best possible solution of SVDP² can be computed within polynomial time. Our strategy consists of the following steps.

- (a) Our key observation on this step is that as we stated in Theorem 1, once a node $v_k^{(i)}$ receives a vaccine, then all nodes $v_k^{(j)}$ with $j \geq i$ and their corresponding edges are removed from \hat{G}' . Based on this observation, we first construct a subset $X_k^{(i)} = \{v_k^{(i)}, v_k^{(i+1)}, \dots, v_k^{(T)}\}$ for each node $v_k^{(i)}$ in $V(\hat{G}') \setminus (\{s, t\} \cup V(G_0))$. Note that $X_k^{(i)}$ in fact is the subset of nodes which should be removed from \hat{G}' once we determined to give a vaccine to $v_k^{(i-1)}$. This takes polynomial time as the number of such subset $X_k^{(i)}$ is equivalent to the size of $V(\hat{G}') \setminus (\{s, t\} \cup V(G_0))$.
- (b) Consider $\mathcal{Q} = \{Q_0 = (p_0 = 0, q_0), Q_1 = (p_1, q_1), \dots, Q_l = (p_l, q_l)\}$. For each Q_i , we are allowed to pick q_i nodes in \hat{G}' after the p_i -th moments and give a vaccine to it, which will eliminate all corresponding nodes (i.e. the nodes in the corresponding $X_k^{(i)}$) from \hat{G}' .

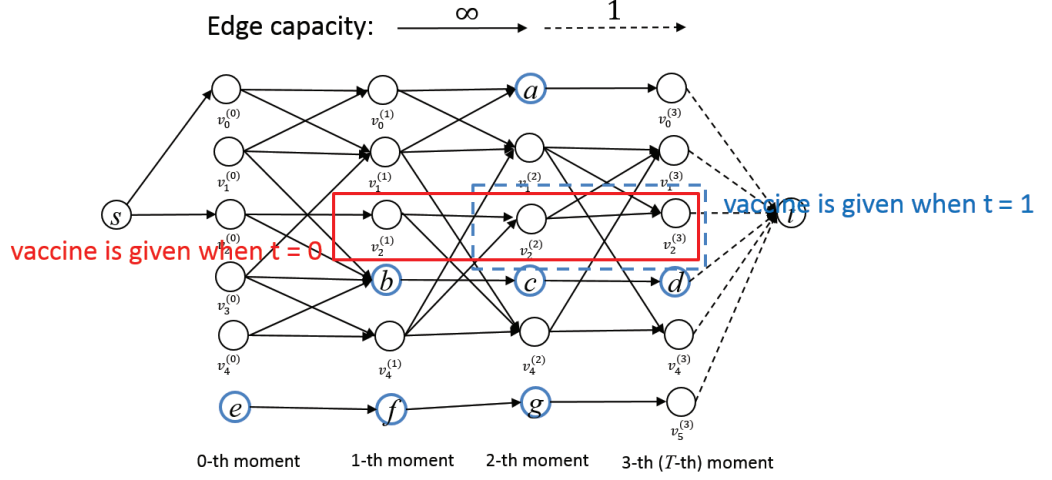


Fig. 3. In this graph, v_2 can be vaccinated when $t = 0$, which will make $X_2^{(0)} = \{v_2^{(1)}, v_2^{(2)}, v_2^{(3)}\}$ removed from the graph, or when $t = 1$, which will make $X_2^{(1)} = \{v_2^{(2)}, v_2^{(3)}\}$ removed from the graph.

Given \mathcal{Q} , the number of all possible choices to select nodes to give a vaccine is bound by $\binom{n}{qT} = O(n^{qT})$, where $q = \max_{1 \leq i \leq T} q_i$. Then, we just need to pick the best one among the all possible choices.

As a result, we obtain a polynomial time exact algorithm for SVDP². Now, we show the correctness of this algorithm.

Theorem 2. *Given T and $q = \max_{1 \leq i \leq T} q_i$ being fixed. The proposed strategy computes the best possible solution within polynomial time.*

Proof. Note all possible choices strategy of giving of vaccines to nodes in \hat{G} is bounded by $O(n^{qT})$. And According to Theorem 1, each time when can we use max-flow algorithm to compute the number of infected nodes, which runs in polynomial time. Thus, the time complexity of our is polynomial under the given assumptions.

5 Concluding Remarks and Future Works

During the recent decade, we have witnessed several new epidemics which has threatened the existence of mankind. In most cases, it took a long time to pro-

duce sufficient amount of effective vaccines, and it took even long to invent a complete cure of it. Therefore, it is of great importance to develop an efficient strategy to minimize the impact of the epidemic while only a limited amount of vaccines are available. This paper aims to open a discuss on this research direction, which is relatively not well understood yet. Our approach uses existing social relationship project strategies to capture the images of evolving social relation which are used to predict the routes of infection of a critical disease. Then, we develop a polynomial time exact algorithm to establish vaccine distribution plan based on the knowledge of future vaccine production and the exacted time to discover a complete cure. We believe that this work shows one significant potential of the information from social network, which is already considered to be with rich set of information for various applications [11–15]. As a future work, we plan to further study the problem to introduce a faster algorithm for it because the running time of our algorithm is very large even though it is polynomial. We are also interested in real data to validate the actual effectiveness of the proposed approach.

References

1. Y. Zhang and B.A. Prakash, “DAVA: Distributing Vaccines over Networks under Prior Information,” *SIAM International Conference on Data Mining (SDM 2014)*, pp. 46-54, 2014.
2. A. Hayrapetyan, D. Kempe, M. Pál, and Z. Svitkina, “Unbalanced Graph Cuts,” *Proceedings of 13th Annual European Symposium on Algorithms (ESA 2005)*, pp. 191-202, 2005.
3. R. Engelberg, J. Könemann, S. Leonardi, and J. Naor, “Cut problems in Graphs with a Budget Constraint,” *Journal of Discrete Algorithms*, vol. 5, issue 2, pp. 262-279, June 2007.
4. S. Finbow and G. MacGillivray, “The Firefighter Problem: A Survey of Results, Directions and Questions,” *Australasian Journal of Combinatorics*, vol. 43, pp. 57-77, 2009.
5. E. Anshelevich, D. Chakrabarty, A. Hate, and C. Swamy, “Approximation Algorithms for the FireFghter Problem: Cuts over Time and Submodularity,” *Proceedings of 20th International Symposium on Algorithms and Computation (ISAAC 2009)*, pp. 974-983, 2009.

6. E. Anshelevich, D. Chakrabarty, A. Hate, and C. Swamy, "Approximability of the Firefighter Problem Computing Cuts over Time," *Algorithmica*, vol. 62, pp. 520536, 2012.
7. N. Wang, "Modeling and Analysis of Massive Social Networks," Ph.D. Dissertation, 2005.
8. B. Bringmann, M. Berlingerio, F. Bonchi, and A. Gionis, "Learning and Predicting the Evolution of Social Networks," *IEEE Intelligent Systems*, vol. 25, issue 4, pp. 26-35, 2010.
9. S.A. Vinterbo, "Privacy: A Machine Learning View," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, issue 8, pp. 939-948, August 2004.
10. V. Kostakos, "Temporal graphs," *Physica A: Statistical Mechanics and its Applications*, vol. 388, issue 6, 1007-1023, 2009.
11. E. Anifantis, E. Stai, V. Karyotis, and S. Papavassiliou, "Exploiting Social Features for Improving Cognitive Radio Infrastructures and Social Services via Combined MRF and Back Pressure Cross-layer Resource Allocation," *Computational Social Networks*, vol. 1, issue 4, 2014.
12. Z. Lu, L. Fan, W. Wu, B. Thuraisingham, and K. Yang, "Efficient Influence Spread Estimation for Influence Maximization under the Linear Threshold Model," *Computational Social Networks*, vol. 1, issue 2, 2014.
13. H. Kim, K. Beznosov, and E. Yoneki, "A Study on the Influential Neighbors to Maximize Information Diffusion in Online Social Networks," *Computational Social Network*, vol. 2, issue 3, 2015.
14. C. Ai, W. Zhong, M. Yan, and F. Gu, "A Partner-matching Framework for Social Activity Communities," *Computational Social Networks*, vol. 1, issue 5, 2014.
15. M. Ventresca and D. Aleman, "Efficiently Identifying Critical Nodes in Large Complex Networks," *Computational Social Networks*, vol. 2, issue 6, 2015.