On Efficient Vaccine Distribution Strategy to Suppress Pandemic using Social Relation

Donghyun Kim

Division of Algorithms and Technologies for Networks Analysis, Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam.
Department of Mathematics and Physics, North Carolina Central University, Durham, NC 27707, USA. donghyun.kim@tdu.edu.vn, donghyun.kim@nccu.edu

Hao Guo
Department of Mathematics and Physics, North Carolina Central University, Durham, NC 27707, USA. hguo@eagles.nccu.edu

Wei Wang

School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an, China. wang_wesw@163.com

Joong-Lyul Lee

Department of Computer Science, University of Texas at Dallas, Richardson, TX 75080, USA. joonglyul.lee@utdallas.edu

Sung-Sik Kwon, Alade O. Tokuta

Department of Mathematics and Physics, North Carolina Central University, Durham, NC 27707, USA. {skwon, atokuta}@nccu.edu

In this paper, we investigate the problem of how to distribute vaccines, which will be supplied over time, so that the number of infected can be minimized during a given mission period. The concept of temporal graph is adopted to abstract the constantly changing social relations over time. Then, we formally introduce the social-relation-based vaccine distribution planning problem (SVDP) on the temporal graph. To solve the problem, we first introduce a new graph induction technique to combine the subgraphs in the temporal graph into a single directed acyclic graph. Then, we design a new technique based on a maximum flow algorithm to evaluate the quality of any feasible solution of the problem. Finally, we propose an enumeration algorithm which will search the solution space using the evaluation technique and find the best possible solution within polynomial time. Our simulation result shows the proposed algorithm is more efficient.

A part of this paper was presented in CSoNet 2015 [1].
This work was supported in part by US NSF No. HRD-1345219.
Corresponding Author. This work was also supported by National NSF of China No. 11471005.
This work was supported in part by US NSF No. HRD-1533653.
than a simple strategy which randomly distributes vaccines.

Keywords: Sensor networks, hybrid sensor networks, barrier-coverage, fault-tolerance, graph theory, optimal solution

1. Introduction

In 2014, our generation has observed the unprecedented spread of Ebola, which is a critical disease and transmittable from an infected person, who has spent a certain incubation period after the initial infection, to another healthy one throughout direct contact to bodily fluid from the infected. Once the seriousness of the disease was recognized by the public, lots of efforts were made to expedite the development of vaccines and cures to slow down the further spread of the disease and save as many lives as possible. As of October 2015, Ebola is largely tamed, mostly thanks to the natural causes. Unfortunately, many people have been deceased during the course of developing cures and vaccines. The bottom line is that 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 significance to save lives, possibly in the very near future.

It is likely that even after a vaccine is invented against a new pandemic, its near-term availability would be highly limited until its mass production. This motivates us to develop a proper plan to distribute vaccines, which are currently available as well as will be available in the future, over healthy individuals because this is as important as inventing new vaccines and cures of unknown pandemic to reduce the number of the victims of a critical epidemic. This social relation information is apparently a great resource to establish a vaccine distribution plan for those epidemics which are spreading from one to another following their (physical) social interaction. However, we found that there is generally a lack of efforts to utilize this approach and this topic needs further investigation.

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. However, as like the other existing works, this approach establishes the vaccine distribution plan only with the currently available vaccines. The vaccines which are supplied over time can be distributed as soon as they are available by using some existing strategy which does not consider the future supply. However, it is not difficult to see that such approach suffers from a bad performance as it is essentially a greedy strategy to achieve the local optima instead of the global optima.

In order to fill this deficiency, this paper introduces a new vaccine distribution algorithm in which vaccines will be supplied over time. Under the assumption that the disease has a very high infection ratio, e.g. 100%, and we can predict the social interaction among the members of the society based on the existing statistics, we exploit the current and the future 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 (SVDP$^2$), which aims to study the best strategy to distribute vaccines which are supplied over time with the objective of minimizing the number of infected until a complete cure is developed. Note that to the best of our knowledge, our algorithm is the first one to compute a distribution plan of both current available and soon-to-be-available vaccines at the same time.

(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$^2$. Then, we propose a polynomial time exact algorithm for SVDP$^2$ by exploiting our evaluation strategy. In addition, we conduct the simulation to evaluate the performance of our algorithm. Our simulation result shows the proposed algorithm outperforms its alternative strategy, which randomly distributes available vaccines.

2. Related Work

Over years, several (computational) vaccine distribution problems were appeared 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. Note that we do not discuss the work in (non-computational) epidemiology or in network security as they are remotely related to the computational vaccine distribution problems of our particular interest.

2.1. Fire-fighter Problem and Its Variations

There are two variations of the the fire-fighter problem [4,5,6]. 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.

2.2. Graph-cut Problems

The main objective of graph-cut problem [2,3] 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.

2.3. Data-Aware Vaccination Problem

Recently, Zhang and Prakash [1] 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 SVDP$^2$. In the following section, we provide the formal definition of SVDP$^2$.

3. Problem Definition

In this paper, we use the concept of the temporal graph [10] $\mathcal{G} = \{G_0 = (V_0, E_0), G_1 = (V_1, E_1), \ldots, G_T = (V_T, E_T)\}$, 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). We assume that $T$ is big enough so that after the final moment, it is highly likely that a complete cure of the disease will be developed. The time gap between two consecutive moments could range from an hour to weeks depending on the context. 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 the existence of the method 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, but we simply assume that the algorithm used for this purpose is highly precise. Now, due to the gravity, let use reiterate our main assumptions in more detail with corresponding justifications if necessary.

(a) The temporal graph $\mathcal{G} = \{G_0 = (V_0, E_0), G_1 = (V_1, E_1), \ldots, G_T = (V_T, E_T)\}$ representing the interaction of the members in the society at each moment is known in advance. 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 $G_{i+1}$. In this paper, we consider the worst-case scenario (in which the infection ratio from two people is 100% if they are related). This means that our approach 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 $\mathcal{Q} = \{Q_0 = (p_0 = 0, q_0), Q_1 = (p_1, q_1), \ldots, Q_l = (p_l, q_l)\}$ are known in advance, where $Q_i = (p_i, q_i) \in Q$ 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.

\textbf{Definition 3.1 (Social-relation-based Vaccine Distribution Planning Problem (SVDP$^2$)). Given $\mathcal{G}, \mathcal{Q}, I_0,$ and $T$, the goal of SVDP$^2$ 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. Proposed Vaccine Distribution Strategy

In this section, we introduce the proposed strategy to distribute the vaccines which are supplied over time. In the first subsection, we discuss how to construct a consolidated graph from the subgraphs in the temporal graph. In the following subsections,
we discuss how to evaluate the quality of a given vaccine distribution strategy and how to find the best vaccine distribution strategy within the polynomial time, respectively.

4.1. Consolidating Subgraphs in Temporal Graph $\mathcal{G}$ into One Integrated Graph $\hat{\mathcal{G}}$

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 $\hat{\mathcal{G}} = (\hat{V}, \hat{E})$, and redefine SVDP$^2$ 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 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:

(i) For each \( v_j^{(i)} \in G_i \) and its neighbor \( v_k^{(i)} \in G_i \), add an direct edge from \( v_j^{(i)} \) to \( v_k^{(i+1)} \) to \( \hat{E} \) only if \( v_k^{(i+1)} \in G_{i+1} \); this means that a node neighboring to an infected node at \( i \)-th moment will be infected in \( i + 1 \)-th moment. Also add a directed edge from \( v_j^{(i)} \in G_i \) to \( v_j^{(i+1)} \in G_{i+1} \) to \( \hat{E} \); this means an infected node at \( i \)-th moment will stay infected in \( i + 1 \)-th moment.

(ii) For each \( v_j^{(i)} \in G_i \) and a virtual node \( w_j^{(i+1)} \), add a direct edge from \( v_j^{(i)} \) to \( w_j^{(i+1)} \) to \( \hat{E} \); this means that once a node is infected, then it will stay infected even though it moved out from the society.

(iii) For each \( v_j^{(i)} \in G_i \) and its neighbor \( v_k^{(i)} \in G_i \), if \( v_k^{(i+1)} \notin G_{i+1} \), then add an direct edge from \( v_j^{(i)} \) to a virtual node \( w_j^{(i+1)} \) to \( \hat{E} \); this means that a node neighboring to an infected node at \( i \)-th moment will be infected in \( i + 1 \)-th moment even though it left the society.

(iv) For each virtual node \( w_j^{(i)} \in G_i \) and a node \( v_j^{(i+1)} \in G_{i+1} \), add a direct edge from \( w_j^{(i)} \) to \( v_j^{(i+1)} \) to \( \hat{E} \); this means that when a node is previously infected, then it is still infected when it returns back to the society.

(v) For each virtual node \( w_j^{(i)} \in G_i \) and a virtual node \( w_j^{(i+1)} \in G_{i+1} \), add a direct edge from \( w_j^{(i)} \) to \( w_j^{(i+1)} \) to \( \hat{E} \); this means that an infected node which left the society is keep infected while it is out of the society.

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 SVDP\(^2\). 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 \( \infty \). Next, for all nodes at the \( T \)-th moment, we add an edge from each of them
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 3-th moments, which is 4 in this example.

to \( t \) with an edge capacity 1 (see Fig. 3). Suppose the resulting graph is \( \hat{G}' \). Then, we prove the following theorem.

**Theorem 4.1.** 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 \( |f| = |I_T| \). \( \square \)
On Efficient Vaccine Distribution Strategy to Suppress Pandemic using Social Relation

4.3. Polynomial Time Exact Algorithm for SVDP² 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 4.1, 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 \( G' \). Based on this observation, we first construct a subset \( X_k^{(i)} = \{v_k^{(i)}, v_k^{(i+1)}, \ldots, v_k^{(T)}\} \) for each node \( v_k^{(i)} \) in \( V(G') \setminus (s,t) \cup V(G_0) \).

Note that \( X_k^{(i)} \) in fact is the subset of nodes which should be removed from \( G' \) 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 \( V(G') \setminus (s,t) \cup V(G_0) \).

(b) Consider \( Q = \{Q_0 = (p_0 = 0, q_0), Q_1 = (p_1, q_1), \ldots, Q_l = (p_l, q_l)\} \). For each \( Q_i \), we are allowed to pick \( q_i \) nodes in \( G' \) 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 \( G' \)).

Given \( Q \), the number of all possible choices to select nodes to give a vaccine
is bound by

\[
\binom{n}{qT} = O(q^{qT}),
\]

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 SVDP^2. Now, we show the correctness of this algorithm.

**Theorem 4.2.** 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 we can 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. Simulation Result and Analysis

In this section, we evaluate the performance of our algorithm via simulation. Note that our problem of interest is new and therefore, there is no existing result directly applicable to solve the problem. Meanwhile, we believe that the random distribution of the vaccines whenever they are available is one reasonable and straight-forward approach to solve this problem. As a result, we compare the performance of our algorithm against the random distribution approach.

For this simulation, we use the social circle data from Stanford Network Analysis Project (SNAP) [11]. In detail, we take one social network graph of 150 nodes (file name: 414.edges) from ego-Facebook dataset in the project SNAP. The social network is visualized in Figure 4. In this simulation, we assume there is initially a single infected personal, and whomever adjacent to an infected personal will be infected if not vaccinated in the following period. For the sake of the simplicity of the simulation, we assume the social relation does not change over time. Then, we observe how the disease spreads for the next four consecutive time periods from the initial status. From the initial time period, we assume there are 1, 4, 7, and 8 vaccines available, respectively. Note that the node who received a vaccine and its adjacent edges will be removed in the next time period.

Figure 4 shows our simulation results. In the figure, the white nodes and the black nodes represent the healthy nodes and infected nodes, respectively. Initial-(0) shows that initially, only one node is infected. Each of RD-(i) for \( i = 1, 2, 3, 4 \) shows how the diseases spreads when the random vaccine distribution strategy is used at each \( i \)-th time period. Meanwhile, each of SVDP-(i) for \( i = 1, 2, 3, 4 \) shows how
Fig. 4. In our simulation, given the same initial infected node, we test the performance of the random vaccine distribution strategy (RD) and our proposed strategy (SVDP). As we can observe from the figure, our strategy can effectively identify the key node connecting two clusters of the people and contain the infection within a single cluster. By comparing RD-(4) and SVDP-(4), we can learn that our strategy can effectively cut the nodes between two clusters so that the infection can be effectively contained.

Figure 5 shows how many people infected in each time period. This figure shows
that our strategy can reduce the number of infected more than 55% less than the random distribution strategy after 4 time period. Our algorithm also suppress the spread of the disease much earlier, i.e. the curve becomes less steep earlier. Apparently, this performance gap will grow as time goes by. From both figures, we can conclude that our strategy has a good potential to save a huge number of lives against a highly contiguous disease like Ebola.

6. 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 produce 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. 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. We also believe the design an efficient distribution plan of both limited cures and limited vaccines would be interesting.
On Efficient Vaccine Distribution Strategy to Suppress Pandemic using Social Relation

References


