Who to vaccinate? A probabilistic view

Guirong Fu*

05 March 2019

1   Background

Assume an infectious disease appear in a known social network, the current infection state is known and only limited vaccinations can be provided. How should we select the vaccinators?

In this article, I try to solve this problem from a probabilistic view. First, I will give a general idea of this model. Then, the specific definitions of each variable are given. Afterwards, I will show how to solve this model and give an application case. Some remarks are made in the last part.

2   Main Idea

To decide which people to vaccinate, we first need to clarify our goal. Why to select these people? Obviously, we want to make the total network's infection rate as low as possible. If we view it from a probabilistic way, that is minimizing the sum of each person's infection probability.

Now, let's considering what influences affect a person's infection probability. It is easy to list the following factors: this person's self-immunity level, the probability of him/her to get touch with other people and the infection state of whom he/she gets touch with. Of course, we can isolate a sick person from a group to avoid the disease broadcasting. But it is not related to vaccination and sometimes unrealistic or unethical. Considering vaccination, what we can do is changing people's immunity level, which exactly matches our question that which people to vaccinate.

If we write the above idea in a mathematical way, the goal is:

$$\min \Sigma_{i \in N} P(i) \tag{1}$$

$$P(i) \sim immu(i), P(i|j), P(j), \forall j \in N \tag{2}$$

---
*Email: fugr0122@gmail.com

- N is the set of all nodes (one node represents a person in a graph representation of a social network).

- $immu(i)$ represents the immunity level of node i.

- $P(i|j)$ is the conditional probability of node i getting infected when the node j's infection state is given.

- $P(j)$ is the probability that node j is infected.

Furthermore, we can have a more general version of this model by adding one more variable k in our goal:

$$\min \Sigma_{i \in N} P^{(k)}(i) \tag{3}$$

k is the times of multiplication the conditional probability matrix from a mathematical view, which can be interpreted as the time span in the application. That is, how long after we expect to minimize the goal. If $k \to \infty$, we will get the final infection probability of our network. However, whether this model has a stable final state is still an open question.

3 Variable Definitions

After having a general impression of this model, let's define each variables formally and find their relations with the given social network.

3.1 Infection Probability of a Node

Before defining the infection probability of node i, we put forward a basic assumption:
According to the definition of probability, $P(i)$ should not be larger than 1..

Definition 3.1.

$$P(i) = \min\{1, immu(i) \cdot \Sigma_{j \in N} P(i|j) P(j)\} \tag{4}$$

3.2 Immunity Level

In this model, we use $immu(i)$ to reflect person i's immunity level, which is the vaccination can change. Of course, the immunity level of a certain disease is determined by many factors, for example the overall healthy state of this person, and whether he/she is vaccinated.

Basically, we require that

- $immu(i)$ should take the value between 0 and 1 (0 and 1 are included).

- the higher the value of $immu(i)$ is, the more likely person i gets infected.

- If a person is vaccinated and he/she hasn't got the disease before vaccination, he/she will be totally free from this disease. That is, $immu(i) = 0$.

We can do many different evaluations for $immu(i)$ which can satisfy these requirements. The simplest version could be:

$$immu(i) = \begin{cases} 0 & \text{i is vaccinated and not infected} \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

3.3 Conditional Probability

We simply assume that the conditional probability of i to catch the disease given j is infected is only determined by the probability of i having contact with j. And the later probability is influenced by some personal attributes of i and j, like their active level, willingness to contact others, and the weight of their link, in the network context.

We generally use $f(i)$ to evaluate the personal attribute of i, and use w_{ij} to represent their link's weight.

Definition 3.2.

$$P(i|j) = g(w_{ij}, f(i), f(j)) \tag{6}$$

g is a map from the tuple variables to a real value between 0 and 1.

$$g : W \times F \times F \to [0, 1]$$

Also, we can have many designs of g, which need to satisfy the following basic requirements.

- $P(i|i) = 1$

- $P(i|j) = 0$, when $i \neq j$ and $w_{ij}=0$.

The second requirement is from another assumption that people have no edges between each other will not get contacted. To be noticed, g can also be simplified to only based on w_{ij} if we don't have any other information to evaluate the personal attribute $f(\cdot)$

We provide some simple examples of g.

Example 3.1.

$$P(i|j) = \begin{cases} 0 & i \neq j \text{ and } w_{ij} = 0 \\ 1 & \text{Otherwise} \end{cases}$$

Example 3.2.

$$P(i|j) = \begin{cases} \frac{\exp(w_{ij})-1}{\exp(w_{ij})} & i \neq j \\ 1 & i = j \end{cases}$$

3

Therefore, using g, we can generate the conditional probability matrix based on the network's adjacent matrix. This is the key step to transfer the network to a probabilistic problem.

$$g(\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \vdots & w_{nn} \end{bmatrix}) \rightarrow \mathbf{CP} \triangleq \begin{bmatrix} p(1|1) & p(1|2) & \cdots & p(1|n) \\ p(2|1) & p(2|2) & \cdots & p(2|n) \\ \vdots & \vdots & \ddots & \vdots \\ p(n|1) & p(n|2) & \cdots & p(n|n) \end{bmatrix}$$

4   Model Solving

To solve this model, we need to provide an initial probability vector $\mathbf{P^{(0)}}$ based on our current knowledge about the infection situation of the given network.

The solution is an evaluation of **immu** vector to minimize our goal. Precisely speaking, it is to select some i and evaluate $immu(i) = 0$ to minimize $\Sigma_{i \in N} P^{(k)}(i)$ of a given k.

$$\mathbf{immu} = \begin{bmatrix} immu(1) \\ immu(2) \\ \vdots \\ immu(n) \end{bmatrix} = \arg\min \Sigma_{i \in N} P^{(k)}(i) \tag{7}$$

$$\mathbf{P}^{(k)} = \min\{\mathbf{1}, \mathbf{CP} \cdot \mathbf{P}^{(k-1)}\}$$

We can also give some restrictions to **immu** vector, such as considering the number of available vaccinations. For example, if we do 0-1 evaluation to the **immu** vector and we only have 2 available vaccinations, there will be a restriction

$$\Sigma_i immu(i) = n - 2$$

n is the number of nodes.

More generally, we can restrict the total number of $immu(i) = 0$ is p, where p is the number of vaccinations we can have, and $immu(i)$ couldn't be 0 if i isn't vaccinated.

5   Application Case

We use the case in the assignment as an example.

1. In this case, we have little knowledge about the initial state except that the person "57" is infected. So, we randomly generate our initial state vector $\mathbf{P^{(0)}}$ and only restrict $P("57")$ to be 1.

2. we use $g(w_{ij}) = \frac{\exp(w_{ij}) - 1}{\exp(w_{ij})}, \forall i \neq j$ to generate the conditional probability matrix $\mathbf{CP}$.

3. Then, we choose do 0-1 evaluation to **immu** vector and restrict we only have 2 available vaccinations, which is expressed by $\Sigma_{i \in N} immu(i) = n-2$.

Based on these information, we use brutal searching method to find the possible solution of **immu**.

When no other information about the initial state vector $\mathbf{P}^{(0)}$, our model are most likely to choose "7", "46", "68" and "77".

If we assume that the 1-step neighbors of "57" are already infected, the model are changed to choose "52", "36" and "12". These results are consistent with the results we get from other network analysis approaches.
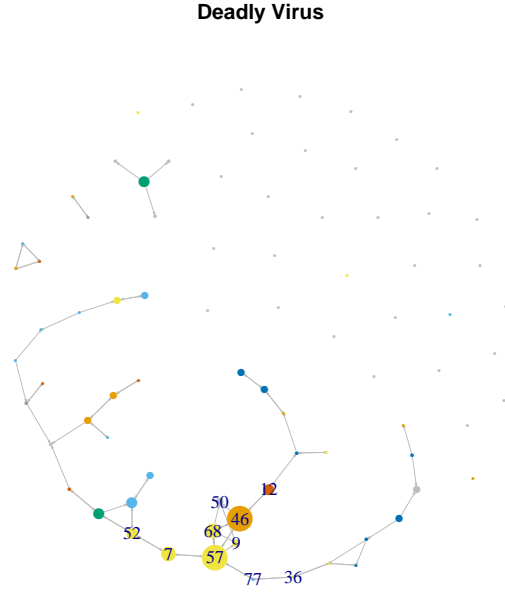
Deadly Virus

Figure 1: The Case: "1057" is infected

6   Some Remarks

This model is very open to allow many different variable definitions based on application context.

However, there are many open questions left:

- How does k influence this model? Does the model have a stable state with k going to infinity?

- Will there always be a solution of **immu**? Will it be unique?

- Any mathematical method to efficiently solve this model?