Optimal Screening Facility Location in An Epidemic

Dingyi Sun
University of Michigan
dysun@umich.edu

Xinyao Zhang
University of Michigan
zxinyao@umich.edu

Zewen Ai
University of Michigan
zewenai@umich.edu

Abstract

The outbreak of coronavirus disease 2019 (COVID-19) has caused a health crisis. In the past three weeks, the virus spread across Washtenaw County, MI. To prevent the county's health system from being overwhelmed, it is necessary to set up some drive-through screening facilities in available pharmacies. For limited test kits and pharmacy locations to launch in facilities, the government have recommended people who contacted with confirmed patients test in nearest medical detection points. Considering the uncertainty of this epidemic, the government would spend less budgets on setting up the test system at this stage and save financial resources to avoid worse situations in the future. Besides controlling budgets, the government had to handle other problems, including transition costs and the cross infection risk. To reduce the spread risk of infectious diseases, people who have test demands should choose the nearest detection point. Through the "linear scalarization" method, risks and costs are included in the total cost with the same unit formulation. To reduce the total cost, we have created an integer programming model to minimize the total cost and solve it through an IP solver under the limited test kits and potential facilities. After participating of volunteers, we could build one more facility and reduce the total cost.

1. Problem Description

In order to let screening facilities being used as soon as possible, we choose existing pharmacy stores to separate space avoiding unnecessary construction. However, different pharmacy locations have diverse available spaces which decide numbers of screening facilities. Medical workers and test resources will be assigned in each facility but limited test kits determine that no more than 200 people will be tested per week in one facility. With participating of volunteers, we can decrease the number of employees then reduce salary expenses, however, we cannot let the facility run with volunteers completely.

In Washtenaw county, residential communities have been divided by ZIP Codes. Residents who are suspected patients or who have closely contacted with patients will become test candidates. The number of test candidates in each community is equal to three times of confirmed cases in this community. After opening facilities, test candidates can complete drive-through tests.

Based on this problem, we create a model to determine the number of opening screening facilities in each pharmacy location as well as the number of weekly tested people in each opening facility. The model should be restricted by minimizing the consumption of test resources, being convenient of residents to get tested, and ensuring a relatively less-crowded test environment to prevent cross infection.

2. Sets, Parameters and Decision Variables

2.1. Sets

$\mathcal{A}$: The set of communities. Communities are classified by zip codes in Washtenaw County[2].
$\mathcal{A} = \{1, 2, 3, ..., 12\}$: $1 \rightarrow 48103$; $2 \rightarrow 48104$; $3 \rightarrow 48105$; $4 \rightarrow 48108$; $5 \rightarrow 48118$; $6 \rightarrow 48130$; $7 \rightarrow 48158$; $8 \rightarrow 48176$; $9 \rightarrow 48189$; $10 \rightarrow 48191$; $11 \rightarrow 48197$; $12 \rightarrow 48198$.

$\mathcal{B}$: The set of pharmacy locations. Pharmacy locations information are represented by zip codes[4].
$\mathcal{B} = \{1, 2, 3, ..., 6\}$: $1 \rightarrow 48103$; $2 \rightarrow 48104$; $3 \rightarrow 48105$; $4 \rightarrow 48118$; $5 \rightarrow 48176$; $6 \rightarrow 48197$.

2.2. Parameters

c: The fixed cost of one screening facility, $c = 30000$. It includes employees' salary, expenses of purchasing test kits and rental fees.

d_{ij}: The traveling distance from community i to pharmacy location j, $i \in \mathcal{A}$; $j \in \mathcal{B}$. Distance data is calculated by Google Map application. A starting point is one of communities and an destination is one of the pharmacy locations, then the distance required to drive is shown in Table 1.

t_i: The test demand in community i, $i \in \mathcal{A}$. Test demands in one community are equal to three times of cur-

Table 1. Driving distance in miles

d_{ij}	j=1	j=2	j=3	j=4	j=5	j=6
i=1	4	8	12	13	8	12
i=2	7	1	6	17	10	6
i=3	9	5	4	19	14	10
i=4	10	5	10	21	5	8
i=5	13	20	23	2	22	25
i=6	7	13	15	9	15	18
i=7	21	24	30	14	16	26
i=8	10	12	18	19	3	14
i=9	11	13	12	18	19	17
i=10	22	15	17	32	14	10
i=11	15	8	11	26	12	3
i=12	14	8	8	24	16	5

rently confirmed cases in this community. From the cases report in Washtenaw County Health Department[1], We conclude that the total population is about three times of lab-confirmed cases, so we refer to this data to determine possible test demands.

Table 2. Test demands in each community

	i=1	i=2	i=3	i=4
t_i	270	201	207	180
	i=5	i=6	i=7	i=8
t_i	87	78	18	168
	i=9	i=10	i=11	i=12
t_i	21	48	663	555

n_j: The size of the pharmaceutical location j, $j \in \mathcal{B}$. This value is proportional to the space of the location and is equal to allowed maximum numbers of facilities.

Table 3. The size of the pharmaceutical location

	j=1	j=2	j=3
n_j	2	2	2
	j=4	j=5	j=6
n_j	3	6	2

w_1, w_2, w_3, w_4: Four weights represent different proportions in the objective.

$w_1 = 0.1$, the maintenance cost takes 1% in the objective. Due to the serious situation, we concern more about the testing ability rather than expenses for opening facilities.

$w_2 = 1$, the traveling cost takes 100% in the objective. In order to let residents arrive facilities with less distance, we assume the traveling cost takes the greatest weights.

$w_3 = 0.2$, the infection risks takes 2% in the objective. The epidemic is highly infectious, so we have to avoid cross infection during testing.

$w_4 = -0.05$, the negative weight with volunteers participates. After volunteers join in facilities, the whole cost will be reduced by less salary expenses.

2.3. Decision Variables

$x_{ij} \in \mathbb{N}$: The weekly tested people from community i to the screen facility j, $i \in \mathcal{A}$, $j \in \mathcal{B}$

$y_i \in \mathbb{N}$: Positive integer indicating how many the screen facility units should open on pharmacy location j, $j \in \mathcal{B}$

$v \in \mathbb{N}$: The number of volunteers.

3. Modeling

We investigated and inspired from a few classical models[6][5][7]. The integer programming model is the formulation which we believe is the best fit to this screen facility arrangement problem, considering the influential factors we listed and the ability and knowledge we have. In order to mathematically solve for the best arrangement, a cost function needs to be presented and minimized. In our model, there are 3 types of cost and 1 more factor been considered. Each of them are separately expressed with linear equation and merged into one cost function.

The first part of the cost is called the maintenance cost. The amount of weekly expenses for supporting a running facility, which should include employee salary, the consumed text kits replenish cost etc. In the previous chapter, we unify the cost for one facility to a constant c. The math formulation for the maintenance cost is showed in Equation 1.

$$\sum_{j \in \mathcal{B}} c y_j \tag{1}$$

The second part of the cost is the traveling distance cost. The sum of the traveling distance of all test candidates. It is impossible to consider the candidates traveling distance individually. Here we generalize the test candidates living location in one zip-code region into a point. Then the traveling distance cost can be easily expressed in a form of double sum in Equation 2.

$$\sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}} d_{i,j} x_{i,j} \tag{2}$$

The last part of the cost is the cross infection cost. We assume the cross infection cost is closely related to the ratio of weekly visiting test candidates in one location to the location capacity in space. In Chapter 2, the simplified location capacity model is given, which is 200 times the size n_j. The formulation for the cross infection cost is given in Equation 3.

$$\sum_{j \in \mathcal{B}} \frac{\sum_{i \in \mathcal{A}} x_{i,j}}{200 n_j} \tag{3}$$

The extra factor we considered is the potential volunteers involved. This factor should be proportionate to the number of volunteers, Equation 4.

$$v \tag{4}$$

The separately modeled costs presents a problem in multi-objective optimization. Solving such a complicated problem is out of our capability. Here we applied the "linear scalarization" method[3], which is commonly used for simplifying and solving multi-objective optimization problems. The key idea is to directly apply weight to each cost function and sum up all parts to give the ultimate objective. In order to intuitively design the weight, the cost functions need to be normalized separately. The idea of "worst case" normalization is applied and map each part of the cost to a number between 0 and 1.

For the maintenance cost, we consider the worst case which we fill up the pharmacy location with facilities. The weekly maintenance cost for such case is:

$$c \times \text{sum}(n) = 510000 \tag{5}$$

Thus we divide this factor with 510000 and times the weight.

For the traveling distance cost, the worst case we take is supposing all test candidates are living one zip-code region and there is only one pharmacy location available with infinite capacity, however, the traveling distance between the zip-code region to the pharmacy location is the longest one. The math formulation goes:

$$\text{sum}(t) \times \max(d) = 74880 \tag{6}$$

For the cross infection cost, we can easily tell the ratio should not exceed 1 and thus for the normalization, we should divide the number of total pharmacy location.

$$\text{size}(\mathcal{B}) = 6 \tag{7}$$

For the volunteer factor, the normalization factor is also trivial. The number of volunteers should not exceed the number of maximum possible number of facility. Thus the factor is:

$$\text{sum}(n) = 17 \tag{8}$$

Fulfilling the model, we add 3 constraints, fully assigned candidates constraint; test capability constraint; location size constraint, and 1 extra constraint limiting the maximum number of total volunteers if volunteers are considered.

The general LP Model without volunteers considered is given:

$$\begin{aligned}
\min \quad & \frac{w_1}{510000}\sum_{j\in\mathcal{B}} cy_j + \frac{w_2}{74880}\sum_{i\in\mathcal{A}}\sum_{j\in\mathcal{B}} d_{i,j}x_{i,j} \\
& + \frac{w_3}{6}\sum_{j\in\mathcal{B}}\frac{\sum_{i\in\mathcal{A}} x_{i,j}}{200n_j} \\
s.t \quad & \sum_{j\in\mathcal{B}} x_{i,j} = t_i \\
& \sum_{i\in\mathcal{A}} x_{i,j} \le 200y_j \\
& y_j \le n_j \\
& x_{ij}, y_i \in \mathbb{N}
\end{aligned}$$

The general LP Model considered volunteers is given:

$$\begin{aligned}
\min \quad & \frac{w_1}{510000}\sum_{j\in\mathcal{B}} cy_j + \frac{w_2}{74880}\sum_{i\in\mathcal{A}}\sum_{j\in\mathcal{B}} d_{i,j}x_{i,j} \\
& + \frac{w_3}{6}\sum_{j\in\mathcal{B}}\frac{\sum_{i\in\mathcal{A}} x_{i,j}}{200n_j} + \frac{w_4}{17}v \\
s.t \quad & \sum_{j\in\mathcal{B}} x_{i,j} = t_i \\
& \sum_{i\in\mathcal{A}} x_{i,j} \le 200y_j \\
& \sum_{j\in\mathcal{B}} y_j \ge v \\
& y_j \le n_j \\
& x_{ij}, y_i \in \mathbb{N}
\end{aligned}$$

4. Results

Results in the model without volunteers include travel transitions and location selections are shown in Table 4. The objective value is 0.5746. Locations in opening facilities are shown in Washtenaw county's map in Figure 1.

Table 4. Results without volunteer

Path	No.	Path	No.	Node	Facility
(2,2)	201	(4,5)	180	y_1	0
(12,2)	112	(8,5)	168	y_2	2
(3,3)	207	(10,5)	48	y_3	2
(9,3)	21	(11,5)	534	y_4	1
(12,3)	172	(11,6)	129	y_5	6
(5,4)	87	(12,6)	271	y_6	2
(6,4)	78	(7,4)	18		
(1,5)	270				

Results in the model with volunteers include travel transitions, location selections and numbers of volunteers are shown in Table 5. The objective value is 0.5715. With participating of volunteers, a new facility can be built in the y_1 location which is the center in the county.

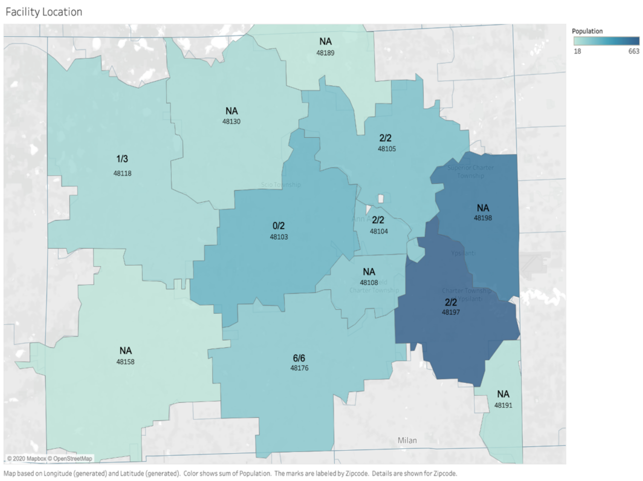

Figure 1. Facilities Location

Table 5. Results with volunteer

Path	No.	Path	No.	Node	Fac.	Vol.	No.
(9,1)	21	(1,5)	270	y_1	1	v	14
(2,2)	201	(4,5)	180	y_2	2		
(12,2)	91	(8,5)	168	y_3	2		
(3,3)	207	(10,5)	48	y_4	1		
(12,3)	193	(11,5)	534	y_5	6		
(5,4)	87	(11,6)	129	y_6	2		
(6,4)	78	(12,6)	271				
(7,4)	18						

For the initial model, we have to build 13 facilities—most of them located in the darker areas on the map which have high test demands. For example, y_6 used up all capacities. The quantities of tests are decided by residential demands nearby. For example, locations of y_2 and y_3 in the center of the county accept many people from high infected areas.

Subsequently, volunteers provide essential assistance. We conclude the new solution. Firstly, a total of 14 volunteers are assigned to different facilities. Also, one more facility is built-in y_1 location. Lastly, the total cost decreased by 0.003.

In summary, where to set up facilities depends on potential test demands. The facilities would be automatically chosen in the center of the area with higher test demands. By increasing volunteers, one more facility would be built and accept test candidates from other regions. It would reduce the cross-infection risk in regions with the high population density and less traveling transitions.

5. Sensitive Analysis

Based on basic results and conclusions, sensitive analyses would be used to test the model to get more insights. We will change the limitation of parameter variables to test the effect on the target variables.

Parameter variables:
1.Infection ratio of candidates
2.Test capacity per facility
3.Full size of facility

Target variables:
1.Object Value
2.Total facility number
3.Facility on each node

The ratio of candidate represents the infectivity of aim disease. It could be used to find the upper-bound of our test system under the limited test kits, the highest cost, and the location selection in each node.

The limited kits affect the upper-bound of testability of our system. From this variable, we can find the change scalar to the total cost under different test capability.

The limited facility affects the upper-bound of testability of our system. From this variable, we can find the change scalar to the total cost under the different full-size facility.

Under the following two variables, it's easy to choose the next step to maintain the system between improvement on test kit ability and facility full-size to keep the lower cost.v

5.1. Infection rate - r

Make the range of r between 1.0 and 4.0, and the first one mentions that the disease has a shallow infection rate as the test candidate is limited to the one person around the confirmed patient. The later one says the high infection rate that the four people around the verified patient should take the kit test.

First check the effect to each node, in the chart below, the dot lines represent the number of facilities in each potential node and increase in a particular order. Node 5 and node six as the area with high confirmed patients will build the facility first, then the rest of the node where launch the facility depends on the distance to the upper infection area in ascending order. If the disease infectivity is low as the ratio at 2, only two nodes should start up the facility, as the infection rate growing, the more node tries to launch the facility.

Turn to the total number of facilities and the object value, cost, it's nearly those total facilities will increase four items when the ratio rises 1 unit. The government has to spend more budget on the test system under a high infection rate. And the cost with ratio 4 is almost twice the one with ratio 2.

5.2. Limited test kits

The number of detection kits reflects the upper limit of the capacity of each detection point, and also limits the

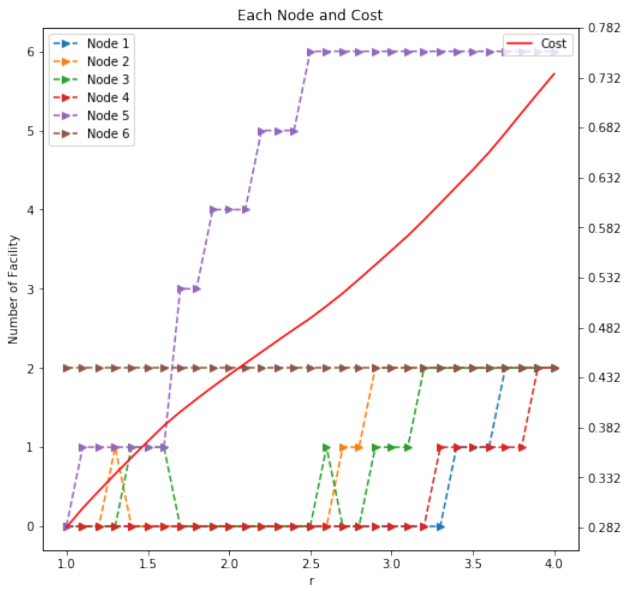

Figure 2. Facility on each node and Cost under different infection ratio

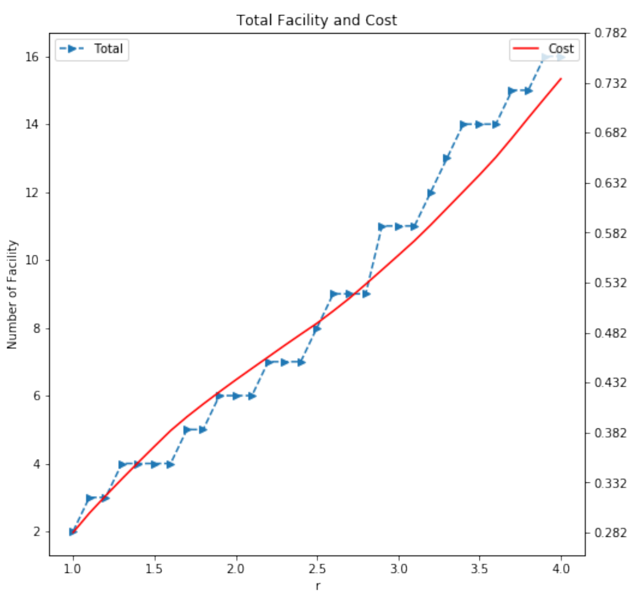

Figure 3. Total Facility and Cost under different infection ratio

maximum detection number of the entire detection system. Logically, if a test point can be completed, the greater the number of tests, the fewer test points we need, but it will increase the risk of infection at the same test point.

First look at the impact of increasing the number of detection boxes on each point. With the improvement of detection capabilities, more and more detection points will be closed. It is interesting that the order of closing these detection points is completely opposite to the order of building. If there are 400 detection kits, only the detection points 5 and 6 will be retained in the end.

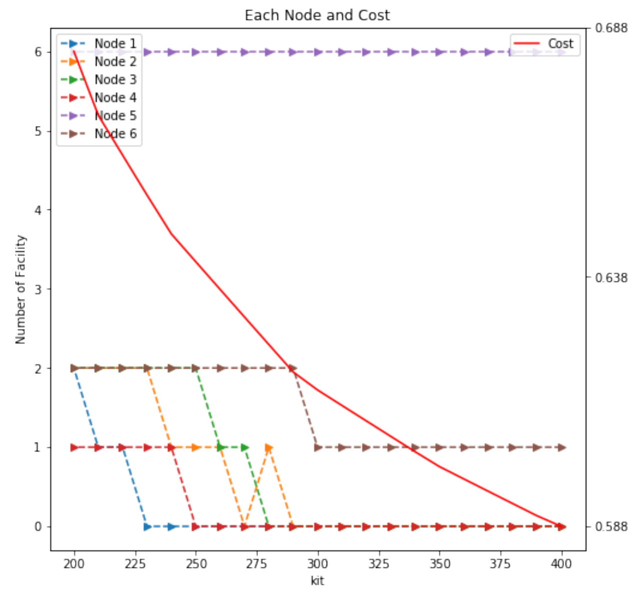

Figure 4. Facility on each node and Cost under different kit

Observing the relationship between the budget required by the government and the number of kits, it can be found that the more kits, the less the budget required by the government, because fewer test points need to be built. Because we give priority to residents who need to take the test to the nearest test site, we believe that short-distance movement is more important than the risk of cross infection.

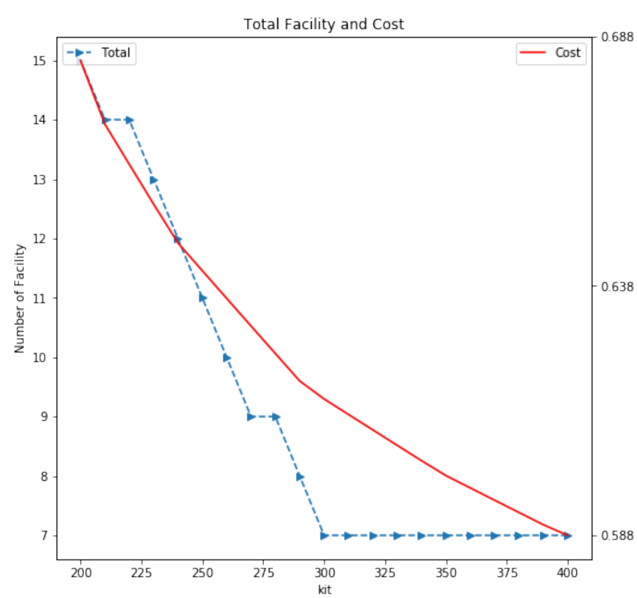

Figure 5. Total Facility and Cost under different kit

5.3. Upper bound of facility in each node

At present, each potential test point has an upper limit for the construction and testing environment. If we limit the number of tests that each facility can perform, and increase the ability of each test point to build facilities, that is, we can build enough facilities now to find new insights.

If you increase the upper limit of the capacity of each detection point, different detection points will have different changes. If the upper limit is increased by only a small amount, areas with high detection requirements will increase the number of facilities, while the remaining areas will decrease. However, as the upper limit of facility increases, the final facilities will gather in nodes 6 and 5 with the highest detection requirements, and node 1 in the county center.

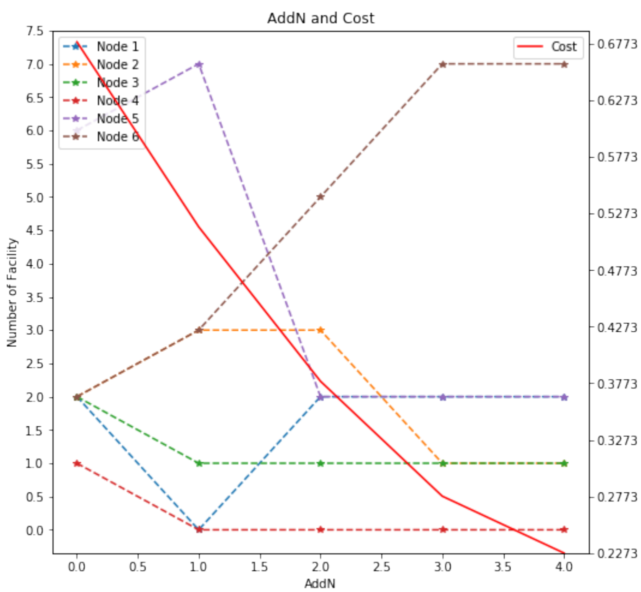

Figure 6. Facility on each node and Cost under increasing size of node facility

As the facility ceiling increases, the total number of facilities that need to be built is decreasing, because the government can set up facilities in the central county and several communities with high demand for testing to reduce expenses. With the decrease in the total number of facilities, the budget required by the government has also decreased significantly. However, based on the currently limited detection capabilities and infection capabilities, at least 13 facilities are also required.

Finally, we compare the impact of increasing the number of kits at each test point (figure 5) and increasing the facility ceiling (figure 7) on government budgets. If the detection capacity is doubled, that is, each facility provides 400 detection boxes, the government budget will be reduced from 0.78 to 0.588. At the same time, if the inspection capacity

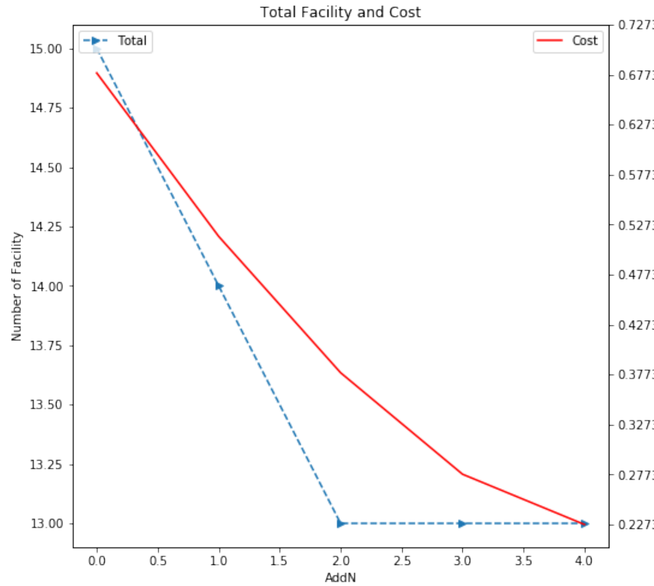

Figure 7. Total Facility and Cost under increasing size of node facility

is not changed, but the construction capacity of 4 additional facilities is provided at a fixed point, the government budget will be reduced from 0.78 to 0.227. The improvement is amazing.

Cross infection risk leads to such a big difference. From the charts of kit and cost (figure 5), we can find that if we do not increase the upper limit of facility and simply increase the number of detection boxes, the total cost will slightly decrease. Although the government saves on spending on building facilities, because the number of facilities is too small, only 6 will guide a large number of residents to the same detection point, which greatly increases the cross infection risk and the possibility of infection. Since our model attaches great importance to cross infection risk, the higher the cross infection risk, the higher the government budget. Compared with the previous charts (figure 5), the facility and cost charts (figure 7) do not increase the number of detections for each facility without increasing the total detection capability. It is easy to find that even if the total of 13 facilities at the end is twice that of the previous sensitivity analysis, the optimized cost suddenly drops to 0.227. This is because in this case, the government reduced the moving distance of most residents and the cross infection risk of the same facility. So guiding residents not to concentrate on the same high-detection facility has a lower cost than inviting them to multiple facilities.

6. Conclusion

Combining the conclusions of the model and the sensitivity test, we draw the following conclusions.

(1) The government should give priority to the establishment of testing points in areas with severe infections and high demand for testing.

(2) When the budget allows, the detection point will expand from high-demand areas to low-demand areas.

(3) Volunteers can effectively help government departments to establish more testing points and reduce budgets under the circumstances of limited medical resources and doctors.

(4) With the increase of the virus infection ability, it is difficult to send all people who have the need for detection to nearby detection points to complete the detection. This medical system will collapse as the ability to transmit infections increases.

(5) Compared to increasing the number of kits for each facility, increasing the number of facilities that can be built by each node will be a more economical method.

(6) For the government, we make the following suggestions:

(i) When the disease is low infectivity ($r < 2$), only the detection points need to be set up near the residents who need more detection.

(ii) As the infectivity of the disease increases, each detection point can gradually establish more facilities and increase the number of points. Facility can be established in high-demand communities first, and then in low-demand communities.

(iii) Over time, the number of detection boxes and detection capabilities will increase. The government should not enhance each facility's testing capabilities. Instead, it should concentrate on county centers and communities with high testing needs to build large testing centers, which is to build as many facilities as possible at each point. Can effectively reduce the cross infection risk and the distance from residents to the detection point, thereby reducing the total government budget.

We are going to split potential patients to the nearest and less crowded pharmacy location to take the test. Our model basically solved the separation and proved advanced analysis. However, disadvantages still exist, like the demand was not dynamic following the general infection model. We could improve that in the future.

References

[1] Cases Report in Washtenaw County Health Department. https://www.washtenaw.org/3108/Cases.

[2] Communities ZIP Codes. https://www.zip-codes.com/county/mi-washtenaw.asp.

[3] Multi-objective optimization Wiki. https://en.wikipedia.org/wiki/Multi-objective_optimization.

[4] Pharmacy Locations Information. https://www.washtenaw.org/BusinessDirectoryII.aspx?lngBusinessCategoryID=22.

[5] David Applegate, William Cook, and André Rohe. Chained lin-kernighan for large traveling salesman problems. *INFORMS Journal on Computing*, 15(1):82–92, 2003.

[6] George L. Nemhauser Gerard Cornuejols and Laurence A. Wolsey. The uncapacitated facility location problem. *Technical Report*, page No. 605, 1984.

[7] Alexander Schrijver. On the history of the transportation and maximum flow problems. *Mathematical Programming*, 91, 2004.