COVID Vaccination Site Selection Problem

Shreyans Jain (202103009)
Aaditya Meher (202103024)
Raj Kariya (202103048)

September 8, 2024

Abstract

Vaccination is the most crucial preventative step during the pandemic. It is inevitable to set up vaccination centers outside of hospitals because the bulk of people will receive vaccinations. Due to the dense population and large volume of daily cases in metropolitan areas, choosing locations for immunization centers presents a significant difficulty for the health sector. Patients, healthcare personnel, the workforce, and the health industry can all suffer from a normal/usual site selection process.

1 Introduction

Coronavirus Disease was first reported in Wuhan, China in March 2019. It was declared as an epidemic in April, 2020. A Country need to distribute the vaccines to safeguard the health of the citizens and its vitality of its economy. To ensure that every individual is vaccinated, this project report presents multiple approaches to select optimal locations from set of existing facilities, with the aim of making the vaccines more accessible and accelerate the recovery of the nation from this pandemic.

This project report concentrates on solving the problem on two different levels, that is, a relatively **small area (Municipal level)** and a relatively **large area (Mass level)**.

2 Mathematical Formulation of the Problem

2.1 Municipal Level

Firstly, let us define some terms:

1. The objective of this problem is to determine optimal location of L (let) vaccination sites from a total of all possible vaccination sites (assume M).

2. Let the municipality be divided into some N number of administrative units (let us call them villages in the further portion of the report).

3. Consider a list of 2-Dimensional vectors, $\{V_i : i = 1, 2, ..., M\}$ which contains locations of all the possible M vaccination sites, which consists of their latitude and longitudes of ith vaccination site.

4. Now, define a similar list of 2-dimensional vectors,$\{B_j : j = 1, 2, ..., N\}$ which consists of the locations of all the N villages, which consists of their latitude and longitudes of jth village.

5. Let the distance between the vaccination site V_i and the B_j village: $d(V_i, B_j)$.

6. Let P_j be the population of jth village and T_p be the total population of the whole municipality. We should also note, that $\sum_{j=1}^{N} P_j = T_p$

For simplicity, let's assume only one vaccination centre for the whole municipality ($L = 1$). So, the optimal distribution strategy should be to allocate the vaccination centres closest to all the B_j and population of the village should also be taken into consideration, that is, vaccination centre should be established where population is maximum. Therefore, we need to evaluate:

$$\min_{1 \leq i \leq M} \sum_{j=1}^{N} \frac{P_j}{T_p} \cdot d(V_i, B_j)$$

Now, to position vaccination sites at more desired locations, we need to consider the factor to locate vaccination sites, near to high number of confirmed COVID patients. To incorporate that, we introduce number of confirmed COVID patients in jth village as C_j and total number of COVID cases as T_c. Therefore, $\sum_{j=1}^{N} C_j = T_c$. Therefore, the updated expression is:

$$\min_{1 \leq i \leq M} \sum_{j=1}^{N} \left[\frac{P_j}{T_p} + \frac{C_j}{T_c} \right] \cdot d(V_i, B_j)$$

At first, we assumed vaccination sites to be only **one**. Let's extend the scenario where the number of vaccination sites denoted by $L \geq 2$ to make it more realistic and general. Our objective is to ensure that the residents of the jth village have access to the nearest possible vaccination site from set of L vaccination sites. The final expression we want to evaluate is:

$$\min_{1 \leq i \leq M} \sum_{j=1}^{N} [\frac{P_j}{T_p} + \frac{C_j}{T_c}] \cdot \min \{d(V_{i_1}, B_j), d(V_{i_2}, B_j), \ldots d(V_{i_L}, B_j)\} - -(1)$$

To make the optimization problem more optimum as a decision support tool, we take into account 2 objectives:

1. Minimize the total distance from the vaccination sites to the villages/administrative units.

2. Maximize the total population within the minimum radius (say ϵ) from the vaccination sites.

2.2 Mass Level

The methodology for mass level vaccination site selection involves a five-step approach to address the real-life problem. The approach involved to model this problem is as follows:

1. The criteria for selecting mass vaccination sites are determined, considering factors such as accessibility, population density, and available infrastructure.

2. Spatial data is collected and mapped using Geographical Information System (GIS) Software to visualize potential locations and their surrounding areas.

3. **Entropy Weighting Method (EWM)** is employed to determine the relative importance levels of the criteria. This method helps in assigning weights to each criterion based on the information provided, ensuring that the most critical factors are prioritized.

4. **Multiple Attribute Utility Theory (MAUT)** approach is used for ranking the potential mass vaccination sites. This technique evaluates and ranks alternatives based on their utility scores, which are calculated by considering the weighted criteria.

5. The ranked alternative sites are analyzed using the network analyst tool of GIS in terms of covered population. This step helps in understanding the effectiveness of each site in reaching the target population and ensuring maximum coverage.

The above approach helps city planners to identify the most suitable locations for mass vaccination sites, ensuring efficient vaccine distribution and maximizing the impact of vaccination efforts.

Mathematically, **Entropy Weighting Method (EWM)** can be formulated with the help of example as follows:

Step 1: Construct the Decision Matrix (D)

$$D = \begin{bmatrix} 5 & 8 & 7 & 9 \\ 6 & 7 & 8 & 6 \\ 9 & 5 & 6 & 7 \end{bmatrix}$$

In this matrix, x_{ij} represents the performance value of alternative i for criterion j.

Step 2: Normalize the Decision Matrix (P)

$$P = \begin{bmatrix} 0.27 & 0.40 & 0.33 & 0.43 \\ 0.33 & 0.33 & 0.40 & 0.29 \\ 0.40 & 0.27 & 0.27 & 0.29 \end{bmatrix}$$

The normalized values (p_{ij}) have been calculated using the formula $\frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}}$.

Step 3: Calculate Entropy Values (E_j)

$$E_j = -\frac{1}{\ln(m)} \sum_{i=1}^{m} p_{ij} \cdot \ln(p_{ij})$$

Using $\ln(3) \approx 1.10$:

$$E_1 \approx 0.35, \quad E_2 \approx 0.50, \quad E_3 \approx 0.38, \quad E_4 \approx 0.45$$

Step 4: Calculate Degree of Diversification (d_j)

$$d_j = 1 - E_j$$

$$d_1 \approx 0.65, \quad d_2 \approx 0.50, \quad d_3 \approx 0.62, \quad d_4 \approx 0.55$$

Step 5: Calculate Criteria Weights (w_j)

$$w_j = \frac{d_j}{\sum_{j=1}^{4} d_j}$$

$$w_1 \approx 0.29, \quad w_2 \approx 0.22, \quad w_3 \approx 0.27, \quad w_4 \approx 0.22$$

These weights (w_j) indicate the relative importance of each criterion.

The above method was the description of the *Entropy Weight Method*. Now., we will take a look at the **Multiple Attribute Utility Theory**, which involves assigning weights to different criteria and scoring alternatives on each criterion to calculate an overall utility score for each alternative.

Step 1: Construct the Decision Matrix (D)

$$D = \begin{bmatrix} 5 & 8 & 7 & 9 \\ 6 & 7 & 8 & 6 \\ 9 & 5 & 6 & 7 \end{bmatrix}$$

In this matrix, x_{ij} represents the performance value of alternative i for criterion j.

Step 2: Assign Weights to Criteria (W)

$$W = [0.2, 0.3, 0.4, 0.1]$$

These weights indicate the importance of each criterion in the decision-making process.

Step 3: Normalize the Decision Matrix (P)

$$P_{ij} = \frac{D_{ij} - \min(D_j)}{\max(D_j) - \min(D_j)}$$

This normalizes the values in the decision matrix to a scale between 0 and 1.

$$P = \begin{bmatrix} 0.333 & 0.667 & 0.5 & 1.0 \\ 0.444 & 0.5 & 0.75 & 0.5 \\ 1.0 & 0.333 & 0.25 & 0.75 \end{bmatrix}$$

Step 4: Calculate Utility Scores (U)

$$U = P \times W$$

$$U = \begin{bmatrix} 0.433 & 0.544 & 0.5 & 0.1 \\ 0.578 & 0.5 & 0.75 & 0.05 \\ 0.2 & 0.1 & 0.1 & 0.075 \end{bmatrix}$$

Step 5: Calculate Overall Utility Score for Each Alternative

$$\text{Overall Utility Score} = \sum_{j=1}^{4} U_j$$

For each alternative:

$$\text{Site 1: } 0.433 + 0.544 + 0.5 + 0.1 = 1.577$$

$$\text{Site 2: } 0.578 + 0.5 + 0.75 + 0.05 = 1.878$$

$$\text{Site 3: } 0.2 + 0.1 + 0.1 + 0.075 = 0.475$$

These overall utility scores help in ranking the alternatives. In this case, Site 2 has the highest overall utility score and may be considered the preferred choice for the mass vaccination site selection problem.

3 Methodology

3.1 Road distance using open street maps

For the ease of overall computation, we utilized some of the python packages, such as OSMNX, which is used to compute the road distance between two points. OSMNX uses OSM data in conjunction with network graphs for a wide range of applications, such as all kinds of urban traffic and planning, all in a network graph analysis framework.

	infected	population	latitude	longitude	Village_name
0	14	2444	13.768463	121.415633	Abung
1	7	2929	13.803248	121.416456	Balagbag
2	2	1933	13.687448	121.438898	Barualte
3	3	2204	13.696362	121.433293	Bataan
4	26	5909	13.790502	121.408104	Buhay Na Sapa
5	3	2492	13.702674	121.400318	Bulsa
6	13	2257	13.825840	121.382144	Calicanto
7	59	5255	13.826772	121.411111	Calitcalit
8	13	2545	13.753426	121.421113	Calubcub I
9	9	4071	13.742145	121.426220	Calubcub II
10	0	1636	13.806595	121.450287	Catmon
11	8	874	13.722816	121.419217	Coloconto
12	8	3755	13.812239	121.364359	Escribano
13	7	1820	13.663068	121.373078	Hugom
14	3	1106	13.690036	121.451034	Imelda (Tubog)
15	1	1743	13.853891	121.367665	Janaojanao
16	14	6927	13.676158	121.387587	Laiya-Aplaya

Figure 1: Table1 : Sample input of Village Centers

	latitude	longitude	Name
0	13.791348	121.407575	San Juan Rural Health Unit II
1	13.826496	121.409996	San Juan Doctors Hospital, Inc.
2	13.824779	121.396247	San Juan Rural Health Unit I
3	13.829028	121.367354	Divine Care Hospital Multi Services Corp.
4	13.809593	121.400683	San Juan District Hospital
...	...	...	...
60	13.831262	121.396933	San Juan Senior High School
61	13.827714	121.394403	Joseph Marello Institute
62	13.829072	121.395489	Batangas Eastern Colleges
63	13.802321	121.403336	Batangas State University - San Juan Campus
64	13.828065	121.394678	San Juan Nepomuceno Church

65 rows × 3 columns

Figure 2: Table2 : Sample input of Vaccination Centers

Figure 3:

3.2 Single-objective Optimization

To numerically solve Equation (1) , we utilized two tables : village centers and vaccination centers.The village centers table includes attributes such as number of COVID-19 patients, population, and geographical location of all villages which is shown in Table-1 .The vaccination centers table contains geographical locations of all villages in the municipality. The vaccination centers is shown in Table-2.

Whenever we are finding the optimal location for a single site optimization problem, we check every possible combination of vaccination centers and villages in each iteration.This approach is called enumeration approach. For smaller values of L($L < 7$), this approach is sufficient enough to give a global minimum. In this approach, checking all possible combinations can be computationally expensive. Therefore, there is a requirement of an efficient and optimized algorithm. Since objective function in Equation (1) is a integer non-linear programming problem, we use Genetic Algorithm for high values of L($L \geq 7$).Genetic Algorithm is a Probabilistic Algorithm which focusses on solution space efficiently by using randomness and probabilistic mechanism.Since it is based on randomness, we end up with different results while running it multiple time.Therefore the algorithm may end up with local solution instead of a global solution.

4 Results

For a sample implementation, we consider an example of San Juan is comprised of 42 villages, whose locations are denoted by $\{B_j, j = 1, 2, 3, \ldots, 42\}$ and we consider 65 vaccination sites in San Juan whose locations are denoted by the $\{V_i, i = 1, 2, 3, \ldots, 65\}$ consisting of 42 elementary schools, 13 junior high schools, 2 senior high schools, 2 universities and 1 church. For optimal location of vaccination sites, we will be selecting optimal sites from one to four vaccination sites that is formulated as $L = 1, 2, 3, 4$. In Figure 1, left figure and right figure illustrate where the best locations are for vaccination sites in San Juan, based on the number of sites.

When there's only one site (L = 1), it's close to the most populated area in the northern part of the town. For two sites (L = 2), one is in the north (yellow star) and the other in the south (purple star). The assigned administrative units are marked with yellow and purple dots accordingly. As we increase the number of sites (L = 3 and 4), the locations become more spread out. Importantly, all optimal sites are along the national highway, meeting the goal of minimizing travel distance from halls.

For L = 4, we observed that the positions of two vaccination sites remained unchanged, while the site in the northern part, which was originally identified as highly populated and with a high number of COVID-19 patients, was replaced by two new sites.

Figure 9 represents a sample data for vaccination centers in 10 villages of San Juan, assuming there are only 2 vaccination sites.

For example the village named 'Balagbag' is linked to vaccination site named 'San Juan Rural Health Unit 1' with distance of 4064.302 meters.Similarly the village named 'Calubcub II' is linked to vaccination site named 'Paaratang Elementary ng Bataan' with distance of 6221.296 meters. We can observe that the village named 'Bataan' with distance of 0 meters to its assigned vaccination site which indicates that the hall of that village and the elementary school of Bataan are in same compund.

5 Conclusion

In this study, we came up with a way to smartly pick locations for COVID-19 vaccination sites in towns. Our method looks at certain parameters such as location of the sites, population density, and number of COVID-19 patients. We made a program that anyone can check to our results.The numerical results show that placing the vaccination sites in specific spots can encourage more people to get vaccinated as early as possible.This method could be useful in poorer countries where there is a lack of public transportation.

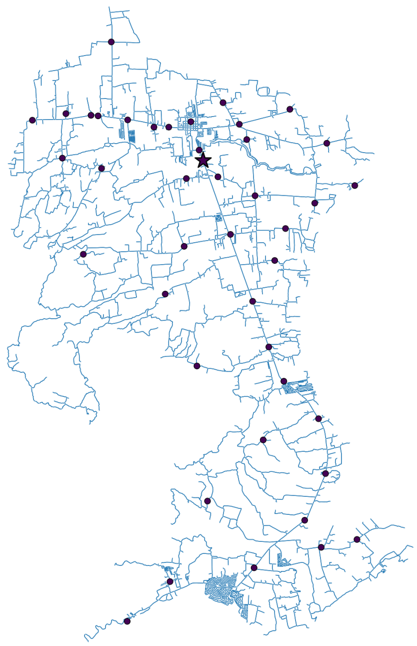

Figure 4: L = 1

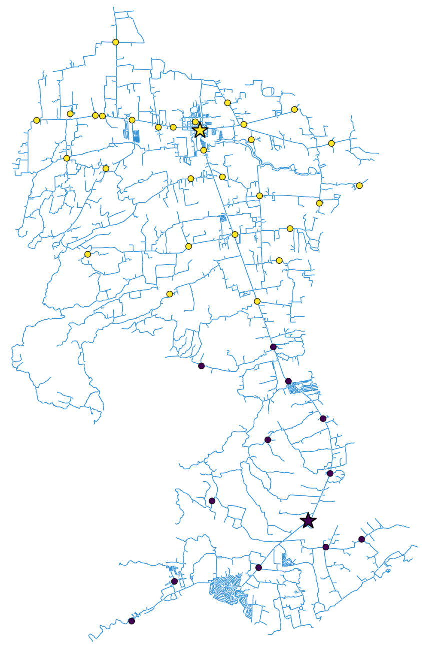

Figure 5: L = 2

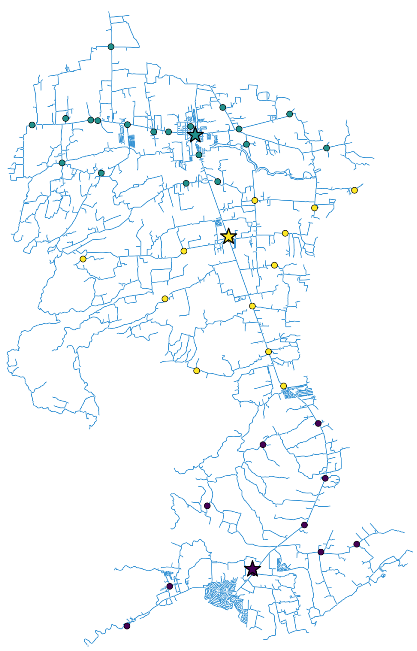

Figure 6: L = 3

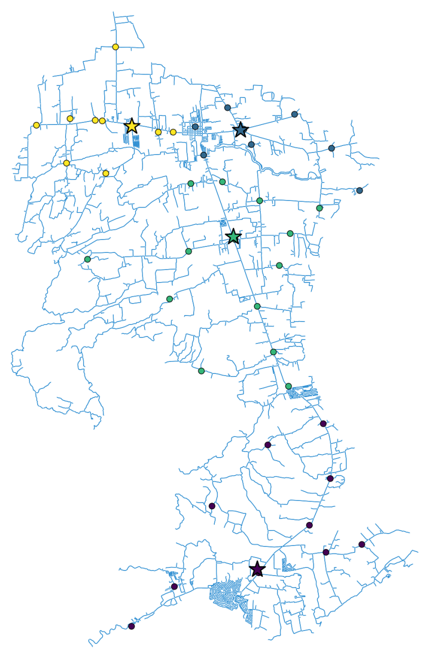

Figure 7: L = 4

Figure 8: Roadmap of San Juan showing the optimal location of one(left), two(right), three(left) and four(right) vaccination sites

Village_name	vaccination_center	distance
Abung	San Juan Rural Health Unit I	6692.135
Balagbag	San Juan Rural Health Unit I	4064.302
Barualte	Paaralang Elementarya ng Bataan	2693.793
Bataan	Paaralang Elementarya ng Bataan	0
Buhay Na Sapa	San Juan Rural Health Unit I	4115.14
Bulsa	Paaralang Elementarya ng Bataan	4552.727
Calicanto	San Juan Rural Health Unit I	1857.728
Calitcalit	San Juan Rural Health Unit I	1684.927
Calubcub I	Paaralang Elementarya ng Bataan	7485.99
Calubcub II	Paaralang Elementarya ng Bataan	6221.296
Catmon	San Juan Rural Health Unit I	7447.483
Coloconto	Paaralang Elementarya ng Bataan	5979.296
Escribano	San Juan Rural Health Unit I	5269.282
Hugom	Paaralang Elementarya ng Bataan	7833.593
Imelda (Tubog)	Paaralang Elementarya ng Bataan	4353.67

Figure 9: Sample Output

Using Genetic Algorithm, we can solve the problem for a greater number of vaccination sites and this method not only apply to municipality but also to big cities and towns. We can apply this method to develop strategies for delivering drug in case of other diseases.

6 References

https://peerj.com/articles/14151/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9212444/pdf/main.pdf

https://colab.research.google.com/drive/1VAh7c3XRqm8qbFu_zFm4ZDQ8GZaO0aT-?usp=sharing