# “Help is on the way!” Optimising Ambulance Locations

Hannah Yasin Ashai
*Department of Electrical Engineering*
*Stanford University*
*Stanford, CA*
*hyashai@stanford.edu*

***Abstract*—Timely access to emergency medical services is crucial in reducing mortality and improving patient outcomes. In many Low and Middle Income countries, long wait times for ambulances pose a significant challenge, often resulting in adverse health consequences. This paper proposes an approach to optimise ambulance locations by formulating the problem using the $p$-Median model, which can be solved using Mixed Integer Linear Programming (MILP) techniques, with the objective of minimising response times. Additionally, we also consider multi-objective optimisation (MOO) case where the aim is to strategically determine the optimal locations for ambulance deployment while also limiting the number of ambulances deployed. This allows us to consider both response times and resource efficiency, both of which are extremely important for LMICs with limited ambulance availability.**

***Index Terms*—ambulance locations, LMIC, Mixed Integer Linear Programming (MILP), multi-objective optimisation (MOO), optimisation, healthcare infrastructure, response times.**

## I. Introduction

The availability and responsiveness of ambulance services play a vital role in saving lives during medical emergencies. However, in Low and Middle Income Countries (LMICs), inadequate ambulance coverage and prolonged response times persist as critical issues, contributing to increased morbidity and mortality rates. This research project aims to address these challenges by leveraging the power of Mixed Integer Linear Programming (MILP) to optimise ambulance locations in LMIC, with a particular emphasis on reducing response times for time-sensitive medical emergencies.

Emergency medical services (EMS) are integral to the functioning of healthcare systems worldwide, regardless of country income levels or geography. They provide immediate medical assistance to patients in emergency situations. Therefore, the design and organisation of EMS systems have a significant impact on patients' outcomes and recovery rates. According to Boutillier and Chan, “Time-sensitive medical emergencies are a major health concern in low and middle income countries (LMICs), comprising one third of all deaths. Examples of such emergencies include cardiac arrest, motor vehicle accidents, and maternal health issues such as childbirth.” [1]

Driven by personal experiences witnessing the repercussions of delayed emergency medical assistance in Pakistan, this project aims to contribute to the enhancement of healthcare infrastructure in LMICs. Insufficient availability and extended wait times for ambulances have often forced individuals to rely on private transportation, which may not provide the necessary medical support during transit.

To effectively address the optimisation of ambulance locations, I adopt a Mixed Integer Linear Programming (MILP) approach to model this problem. Note, the distinction between an MILP and a generic LP is that the decision variables can take on *only* integer values in some component of the problem– such as using a binary variable to reflect the case where an ambulance is placed at a location or not. By solving the MILP model, we can obtain an optimal solution that guides the placement of ambulances in locations that minimise response times. Note that MILP problems are known to be NP-hard, adding to the complexity of finding optimal solutions.

In the context of this MILP approach, addressing resource constraints in LMICs is crucial. For example, Karachi, the largest city in Pakistan with a population exceeding 20 million, has 150-500 ambulances [2] [3], which mostly belong to Non-Governmental Organisations (NGOs), while New York City, with a population a little more than one-third the size of Karachi, has 675 ambulances [4] . This disparity underscores the variations in access to emergency medical resources. To tackle this challenge for LMICs, an additional objective is included: limiting ambulance deployment to optimise resource utilisation. This objective aims to achieve a balance between response times and efficient allocation of available resources.

### *Outline*

In this paper, I begin by reviewing and analysing existing works in the field of ambulance location optimisation and the various models used. Building upon the insights gained from this literature review, I proceed to formulate our problem as a MILP model and analyse the single objective case, where we aim to simply minimise time taken to reach a patient. Then, in order to tackle the challenge of minimising the number of ambulances needed while optimising response times, I employ a multi-objective optimisation (MOO) approach using the weighted sum method with both objectives being given equal weights. Through a comparison with a random placement of the ambulances, I assess the effectiveness of the proposed approaches. Furthermore, I discuss the implications of my findings, potential limitations of my current work, and highlight potential areas for future research

## II. Prior Work

Initially, I referred to large scale literature reviews in the subject domain in order familiarise myself with EMS dispatch protocol in the status quo. In Brotcorne, Laporte and Semet's work, they describe the details of ambulance dispatch; ambulance intervention in emergencies involves incident detection, call screening, vehicle dispatching, and paramedic intervention. Emergency Controllers assess incident severity and urgency for timely ambulance dispatch, strategically placing vehicles to minimise response times and provide prompt medical assistance, enhancing emergency medical services. These ambulance locations are modelled as graphs with a set of demand nodes ($V$) and possible ambulance supply nodes ($W$). Ambulance location models are defined on graphs. The shortest travel time $t_{ij}$ from vertex $i \rightarrow j$ is known. [5]

Regarding the capacity size of ambulances, most literature considers ambulances to be of a set carrying capacity (usually 1 patient) [5], which I used as an assumption in my work as well. I also included coverage constraints to ensure that each demand point is covered at least once (such as in [6]).

With this general structure, there are various formulations of this EMS Location problem. According to Li, Zhu and Wyatt, there are three broad categories of models [7]:

1) Covering Models: These models emphasise providing coverage for emergency calls within a predefined distance standard. The objective is to identify the optimal locations for ambulances such that the coverage requirements are met efficiently.
2) p-Median Models: Such models minimise the total service distance for all demand nodes. The objective is to determine the optimal locations for a fixed number of ambulance ($p$) in order to minimise the overall travel distance or time required to reach all demand points.
3) p-Centre Models: These models aim to minimise the maximum service distance for all demand points. The objective is to identify the optimal locations for a fixed number of EMS facilities ($p$) such that the maximum distance travelled by any demand point to its nearest facility is minimised. These models ensure equitable access to EMS resources.

It should be noted that most EMS models tend to use covering models. [7]. However, I chose to focus on the $p$-Median problem, considering its relevance in emergency services in the context of LMIC. This is because the challenges faced by emergency response systems in the developing world highlight the need for efficient resource allocation and location planning [1]. In particular, by minimising service distances for all demand nodes, the $p$-Median problem offers a valuable approach to optimising the placement of EMS facilities in resource-constrained settings.

Many papers also exist on Multi-Objective Facility Location Problems (MO-FLPs), specifically in the context of ambulance and emergency services. These studies recognise the need to simultaneously consider multiple objectives, such as minimising response times, maximising coverage, minimising costs, and optimising ambulance allocation. [8] I consider the case of adding a second objective function minimising the number of total ambulances allocated.

## III. Problem Formulation

### A. Single Objective Optimisation

I have formulated this problem first as an instantiation, with necessary modifications, of the $p$-Median problem.

TABLE I
NOMENCLATURE

| Symbol | Description |
|---|---|
| $n$ | Total number of patients |
| $p$ | Total number of EMS Ambulance facilities |
| $m$ | Total number of possible locations |
| $t_{ij}$ | Time taken for ambulance $j$ to reach patient $i$ |
| $x_{ij}$ | Decision variable for patient $i$ served by ambulance $j$ |
| $a_j$ | Indicator variable for ambulance location $j$ |

Note that I assume each ambulance has capacity 1.

Our objective function is as follows:

$$\min \sum_{j=1}^{m} \sum_{i=1}^{n} t_{ij} \cdot x_{ij}$$

Such that the following constraints hold:

1) $\sum_{j=1}^{m} x_{ij} = 1 \quad \forall \quad i = 1, \cdots, n$
2) $x_{ij} \in \{0, 1\} \quad \forall \quad i = 1, \cdots, n, \quad \forall \quad j = i = 1, \cdots, p$
3) $\sum_{i=1}^{n} x_{ij} \leq a_j \quad \forall \quad j = 1, \cdots, m$
4) $a_j \in \{0, 1\} \quad \forall \quad j = 1, \cdots, m$
5) $\sum_{j=1}^{m} a_j \leq p$

The objective is to minimise the total time taken to reach all patients by assigning them to appropriate ambulances. The first and second constraint make sure that patient $i$ is provided for by ambulance $j$ and that an ambulance $j$ is either assigned to them ($x_{ij} = 1$) or is not ($x_{ij} = 0$). The third constraint forces the fact if a patient $i$ is assigned to ambulance $j$ that ambulance must be chosen from our set of $m$ location. The fourth constraint ensures an ambulance location $j$ is either chosen 1 or not 0. The fifth constraint is meant to enforce that no more than $p$ ambulances are located. Note, implicitly that $m \geq p \geq n$.

### B. Multi-Objective Optimisation

Our decision variables are the same as before. In addition, our objective functions are now:

- $\min \sum_{i=1}^{n} \sum_{j=1}^{p} t_{ij} \cdot x_{ij}$
- $\min \sum_{j=1}^{m} a_j$

The constraints are the exact same as before for the single objective case! Our only modification is that we seek to optimise two objective functions.

Ultimately, the erighted sum method MOO MILP approach transforms our list of objective functions into the following:

$$\min w_1 \sum_{i=1}^{n} \sum_{j=1}^{p} t_{ij} \cdot x_{ij} + w_2 \sum_{j=1}^{m} a_j$$

where $w_1 = w_2 = 0.5$.

### C. Data Generation

I use synthetic data generated in the following way:

- Each element of our time matrix $t$ denotes the number of minutes it takes to reach patient $i$ from ambulance $j$. I used an uniform random sampling from 1 to 60 to assign a value to each element of our time matrix.
- I trialled different values of $n, m$ and $p$.

## IV. Results

I compared the objective values obtained using two different optimisation libraries: CVXPY and PuLP. I found that the objective values obtained using PuLP were consistently lower, and hence better, than those obtained using CVXPY. Therefore, for clarity's sake, I performed most of my analysis using PuLP.

To solve Multi-objective problems, PuLP uses a weighted sum method, which uses a vector of weights $\mathbf{W}$ to convert the objectives functions $\mathbf{O(x)}$ to a single objective $f$, where elements of $\mathbf{W}$ are non-negative and sum to 1.

$$f(\mathbf{x}) = \mathbf{W}^T \mathbf{O(x)}$$

### A. Single Objective: Fixed $n, p$ and $m$

As a first analysis, I examined the behaviour of our MILP solution in comparison to a random solution by plotting the objective values for various seed values. Due to computational constraints, I used relatively small values for $n = p = 25$, and $m = 35$.

It is evident from the plots that the MILP solution consistently outperforms the random solution by a factor of $10^2$. The objective values obtained from the MILP approach are significantly lower than those from the random approach across different seed values from 0 to 20. This observation highlights the superiority of the MILP method in finding optimal or near-optimal solutions compared to relying solely on randomisation. The substantial performance gap is to be expected in such a case with relatively small values of $n, p$ and $m$. (Refer to Figure 1.)

To check that the performance gap persists for large values of $n, p$ and $m$, I also ran two random instances (for seed values of 0 and 1) of my code for $n = p = m = 1000$. In both cases the minimum objective found by the MILP was 1000, while the random objective was 14966933 and 15004441, suggesting that the divergence between the random solution vs the MILP solution remains very large.

### B. Single Objective: Contour Plots for Fixed $n$

The next mode of analysis I employed involved generating contour plots to explore the difference between the random and MILP objective values, using fixed values of $n$. These contour plots provide a visual representation of the difference between the objective values found for different combinations of $m$ and $p$.

I generated contour plots for different values of $n$ (specifically, $n = 25, 100, 250$) to analyse the difference between the random and MILP objective values for various patient burdens.

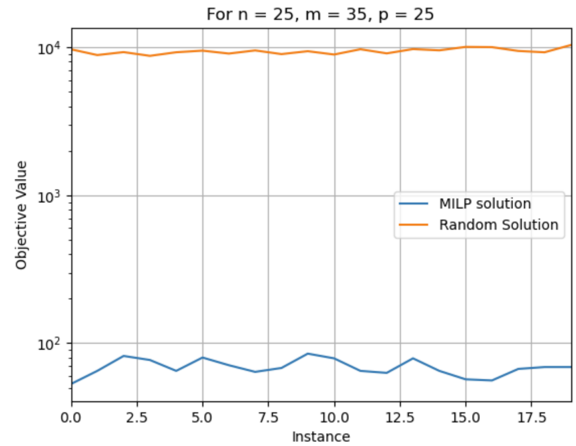


Fig. 1. This figure shows the plots of the objective values of the random and MILP solution for different instances of an initial seed for $n = 25, m = 35$, and $p = 25$. The objective values are plotted on a logscale.

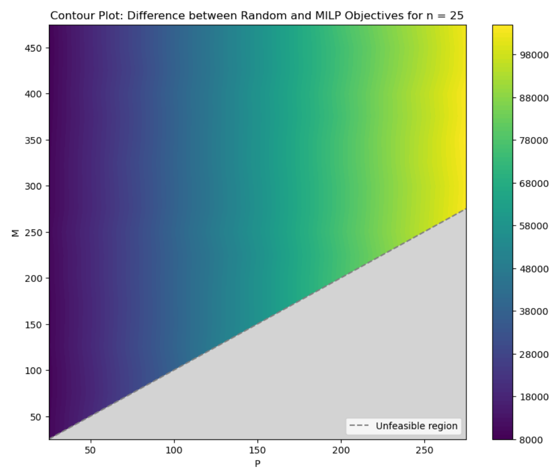


Fig. 2. Contour plot where the colour bar indicates difference between random and MILP objective value for $n = 25$ where $25 \leq m \leq 500$ and $25 \leq p \leq 300$. The grey region is unfeasible.

The contour plots were created by varying the values of $m$ and $p$, where $n \leq m \leq 500$ and $n \leq p \leq 300$. The gray region in the plots represents the unfeasible area where there are no valid solutions.

There are certain similarities across all the different values of $n$. One notable trend is that as the number of potential ambulances increases so to does the difference between the random and MILP objectives. This may indicate that the problem is relatively insensitive to the size of $m$ as compared to the value of $p$.

Interestingly, even when the number of ambulances and facilities is small, there is still a noticeably large between the random and MILP objectives. This gap only grows as the number of patients increases, suggesting that the MILP approach consistently provides superior solutions compared to random selection, regardless of the problem size. For very populated areas, this suggests that applying any kind of MILP

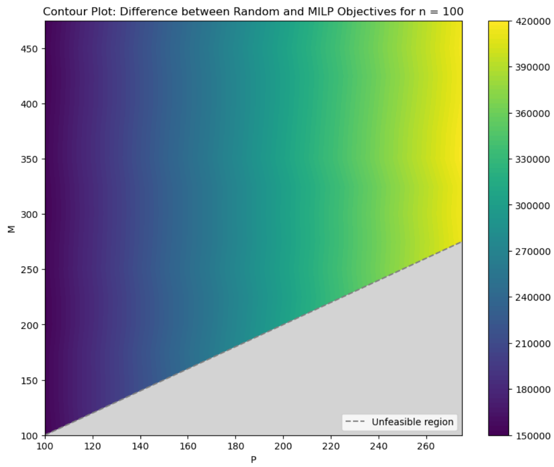


Fig. 3. Contour plot where the colour bar indicates difference between random and MILP objective value for $n = 100$ where $100 \leq m \leq 500$ and $100 \leq p \leq 300$. The grey region is unfeasible.

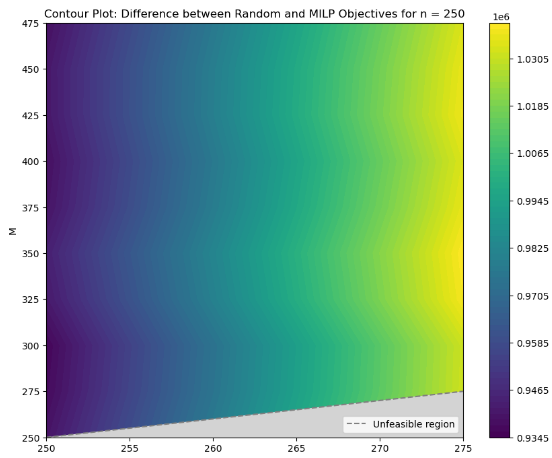


Fig. 4. Contour plot where the colour bar indicates difference between random and MILP objective value for $n = 250$ where $250 \leq m \leq 500$ and $250 \leq p \leq 500$. The grey region is unfeasible.

technique has the potential to greatly reduce response times.

In addition to the observed differences in objective values, another point of note in the contour plots is the increasing 'waviness' of the contour lines as the value of $n$ increases (compare Figure 4 with 3 and 2). This phenomenon indicates a higher degree of variability or fluctuations in the objective difference between the random and MILP solutions. One possible explanation of this is the increased complexity of the problem as the number of patients ($n$) increases; the interactions between ambulance locations and patient assignments become more intricate.

### C. Multi-Objective: Fixed $n, p$ and $m$

The results from the multi-objective case align with our expectations based on the single-objective optimisation for $n = p = 25$ and $m = 35$. Figure 5 demonstrates the significant improvement achieved by the MOO MILP solution compared to a random solution for the first objective. We also see that the second objective function is much lower for the MOO MILP than the random case, although the improvement is on a lower order of magnitude compared to the first objective.

Let us recall here that the first objective refers to the time minimisation problem and the second objective refers to the ambulance minimisation.

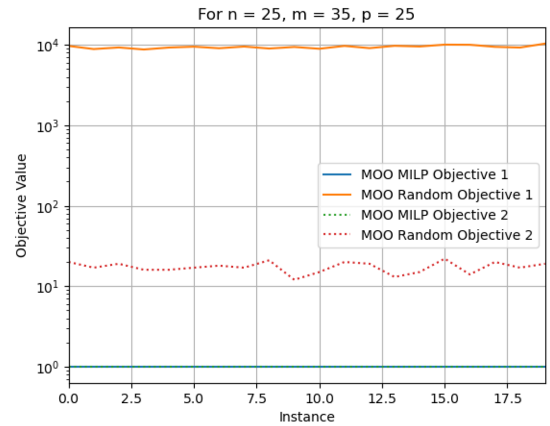


Fig. 5. This figure shows the plots of the objective values of the random and MILP solution for the MOO optimisation different instances of an initial seed for $n = 25, m = 35$, and $p = 25$. The objective values are plotted on a logscale.

### D. Multi-Objective: Contour Plots for Fixed $n$

As before, the next part of my analysis of the MOO MILP involved generating contour plots to explore the difference between the random and MILP objective values for both the first and second objective, using fixed values of $n$. These contour plots provide a visual representation of the difference between the objective values found for different combinations of $m$ and $p$.

I generated contour plots for different values of $n$, specifically, $n = 25, 250$ to analyse the impact of increasing patient burdens. This is the same analysis employed in the single-objective case. The contour plots were created by varying the values of $m$ and $p$, where $n \leq m \leq 500$ and $n \leq p \leq 300$. The gray region in the plots represents the unfeasible area where there are no valid solutions. As in the single objective case, we observe strange waviness that arises for $n = 250$ that is not present for smaller $n$ (see Figure 6 vs Figure 7) and the same relative insensitivity to the size of $m$.

In the case of the second objective, we observe the opposite sensitivity; The strong vertical gradient indicates that larger values of $m$ lead to larger objective values. (See Figure 8 and Figure 9). This observation is logical because increasing the number of possible ambulance locations increases the flexibility in ambulance deployment. With more potential locations available, it is likely that a larger number of ambulances will be needed to cover those locations effectively.

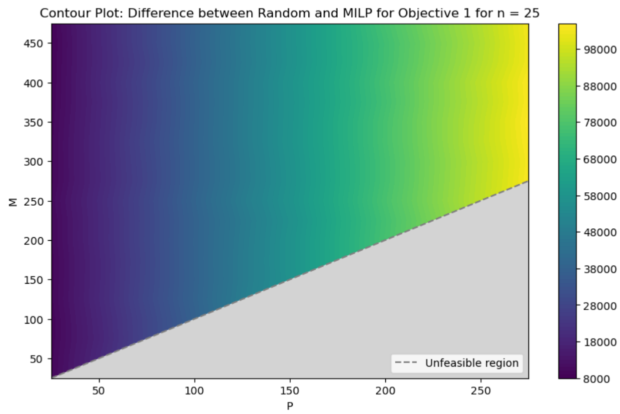


Fig. 6. Contour plot where the colour bar indicates difference between random and MILP for the *first* objective value (time minimisation) for $n = 25$ where $m$ ranges from $n$ to 500 and $p$ from $n$ to 300

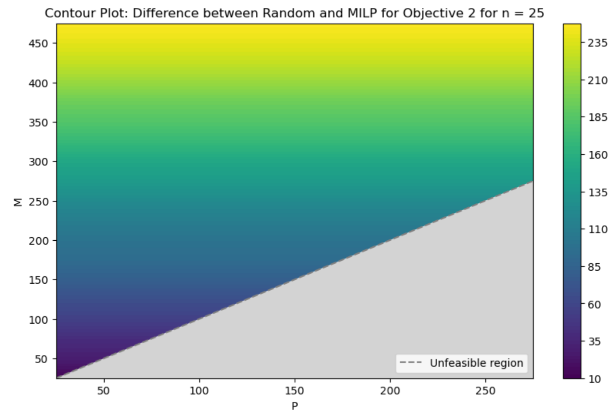


Fig. 8. Contour plot where the colour bar indicates difference between random and MILP for the *second* objective value (resource minimisation) for $n = 25$ where $m$ ranges from $n$ to 500 and $p$ from $n$ to 300

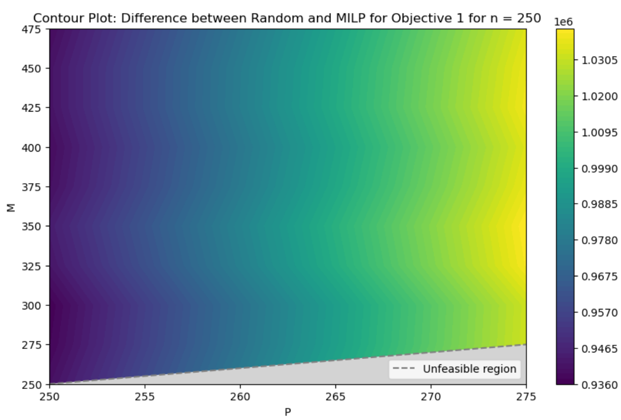


Fig. 7. Contour plot where the colour bar indicates difference between random and MILP for the *first* objective value (time minimisation) for $n = 250$ where $m$ ranges from $n$ to 500 and $p$ from $n$ to 300

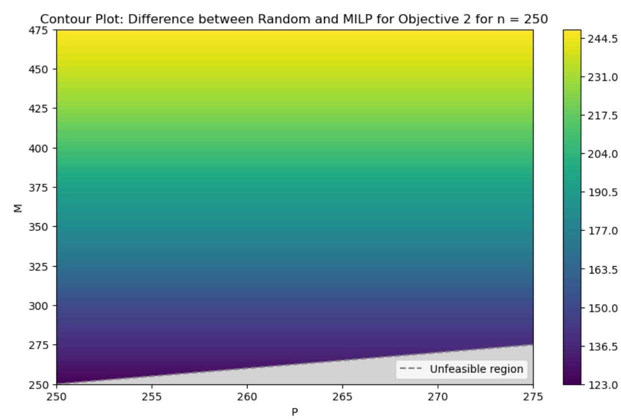


Fig. 9. Contour plot where the colour bar indicates difference between random and MILP for the *second* objective value (resource minimisation) for $n = 250$ where $m$ ranges from $n$ to 500 and $p$ from $n$ to 300

## V. Conclusion and Further Directions

The optimisation of ambulance locations in LMICs is a critical task to ensure timely access to emergency medical services and improve patient outcomes. By leveraging the $p$-Median approach and multi-objective optimisation techniques, we have demonstrated the ability to determine optimal ambulance deployment locations while considering response times and also resource efficiency.

Our results strongly indicate that the MILP approach outperforms random solutions, implying that such a model, if deployed correctly, would be successful in reducing morbidity. However, further enhancements can be made to the methodology, such as through using online methods, integrating probabilistic fitting methods and/or coverage methods to ensure equitable access and better response predictions.

Using online methods, which continuously adapt ambulance locations based on real-time data and demand patterns, can better respond to dynamic changes in emergency service requirements. By updating ambulance deployment strategies in real-time, online methods can improve responsiveness and reduce overall response times.

Additionally, incorporating probability and coverage control mechanisms can account for the uncertainty and variability of emergency incidents. By integrating probabilistic models and considering coverage constraints, ambulance locations can be optimised to ensure sufficient coverage in high-demand areas while maintaining an appropriate level of redundancy.

A caveat of this current project is that I relied on synthetically generated data. Future research should focus on using non-synthetic data to validate and refine the optimisation models. With real-world data on emergency incidents, geographic constraints, and transportation networks, the accuracy of ambulance location optimisation can be further improved.

Overall, the combination of MILP, multi-objective optimisation, online methods, probabilistic fitting and coverage control holds great potential for enhancing ambulance location decision-making in EMS for LMICs. By continuously refining these methodologies, we can strive for more effective emergency medical services that saves more lives.

## APPENDIX

The code repository for this project is included here as a github link: https://github.com/hya00/CS-361-Final-Project

## REFERENCES


[1] Boutilier, J. J., and Chan, T. C. (2020). Ambulance emergency response optimization in LMIC countries. Operations Research, 68(5), 1315-1334.

[2] S. Shackle, ”On the frontline with Karachi’s ambulance drivers,” The Guardian, Apr. 6, 2017. [Online]. Available: https://www.theguardian.com/world/2017/apr/06/on-the-frontline-with-karachis-ambulance-drivers. Accessed June 1, 2023.

[3] T. Ahmed, ”Karachi’s struggle for an adequate ambulance service,” The Express Tribune, May 12, 2022. [Online]. Available: https://tribune.com.pk/story/2356225/karachis-struggle-for-an-adequate-ambulance-service. Accessed June 1, 2023.

[4] Mayor’s Office of Operations. ”Fleet Report.” NYC.gov. https://www.nyc.gov/site/operations/performance/fleet-report.page Accessed June 1, 2023.

[5] Brotcorne, L., Laporte, G. and Semet, F. (2003). Ambulance location and relocation models. European journal of operational research, 147(3), 451-463.

[6] Daskin, M.S. and Stern, E.H. (1981). A hierarchical objective set covering model for emergency medical service vehicle deployment. Transportation Science, 15(2), 137-152.

[7] Li, X., Zhao, Z., Zhu, X. and Wyatt, T. (2011). Covering models and optimization techniques for emergency response facility location and planning: a review. Mathematical Methods of Operations Research, 74, 281-310.

[8] Karatas, M. and Yakıcı, E. (2018). An iterative solution approach to a multi-objective facility location problem. Applied Soft Computing, 62, 272-287.