Optimization Modeling of EMS Dispatch Operations in response to COVID19

Jing Liu, Ruilin Ouyang, Chun-An Chou*

Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, USA

Abstract

In this paper, we present optimization models for determining optimal Emergency Medical Services (EMS) resource allocation or reallocation in response to the demand disruption due to COVID-19. We explore the ambulance trips between boroughs and ambulance dispatch areas during the pandemic period based on a real dataset of Emergency Response Incidents in New York City. From mid-March to mid-April 2020, the demand for ambulances in the five boroughs of New York reached its highest level in the last five years. The ambulance demand increased significantly and the response time performance was starting to be eclipsed. To minimize the gap between ambulance demand and supply, the project develops a dynamic optimization model to identify optimal ambulance resources allocation and dispatch policies updated in real time with minimal cost. The model is able to reallocate the EMS resource with minimum costs when the demand changes unexpectedly while guaranteeing fairness when the demand exceeds the available resource.

Keywords: Integer Programming, EMS

1. Introduction

2. Related Work

3. Problem Definition

Our data is obtained from the New York open data, which contains all EMS activation incidents from 2015 to August 2020, including incident start-time, incident type, incident travel time, incident response time, total incident time, ambulance departure area, and ambulance arrival area, etc. We can classify the ambulance dispatch into three categories: same-area dispatch, cross-area dispatch, and backup area dispatch. Among them, same-area dispatch is the primary dispatch method. Besides, based on our assumptions, there are three possible ways to determine the length of travel: based on the median duration of actual travel time in the near past (January to May 2020); based on the median duration

*Corresponding author

Email address: ch.chou@northeastern.edu (Chun-An Chou)

of real travel time in the last year (the whole year of 2019); actual travel distance estimated from zip code. Our model will compare these three different travel time measurements to see which one should be most close to reality.

According to New York Ambulance data from March to April 2020, COVID-19 led to a significant increase in ambulance demand during those two months and reached a five-year peak. This has led to a significant increase in ambulance response times, patient wait times, i.e.. We aimed to make the emergency medical system able to prearrange the ambulance in advance to alleviate the excessive ambulance response time in case the demand exceeds the supply. Given the daily demand of EMS, availability of the trip, travel time, and the maximal amount of EMS that each dispatch area can be dispatched per day, we established a linear programming model that can minimize the total number of EMS assigned with minimal travel time. Thus, we intended to provide the optimal ambulance allocation strategy.

4. Data Background

We used the EMS Data in New York from 2015 to 2020. There are 5 time points shown in Figure 1: t_1: time of the incident was created; t_2: time of the first ambulance is assigned; t_3: time of the ambulance arrived at the location of the incident; t_4: time of the ambulance arrived at the hospital; t_5: time of the incident closes. Then, there are four time-variables: dispatch response time (between t_1 and t_2), travel time (between t_2 and t_3), incident response time (between t_1 and t_3), and total time (between t_1 and t_5). Incident response time and dispatch response time have highly correlation coefficient of 0.8, while other time variables have no significant correlation. From Figure 2, we can see that the demand for ambulances from mid-March to mid-April 2020 reached its highest level in the last five years. The ambulance demand, average incident response time, average dispatch response time, average total time all increased significantly in these two months, especially from Mar. 22nd to April 8th.

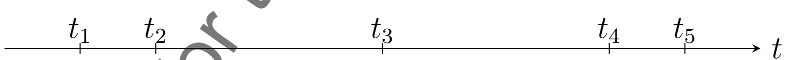


Figure 1: Time Point

In New York, there are 5 boroughs where the incident occurred {Brooklyn, Bronx, Manhattan, Queens, Staten Island}, and 35 ambulance dispatch areas where the ambulance assigned: 7 in the Brooklyn area $\{k_1, K_2, K_3, K_4, K_5, K_6, K_7\}$; 5 in the Bronx $\{B_1, B_2, B_3, B_4, B_5\}$; 9 in the Manhattan $\{M_1, M_2, M_3, M_4, M_5, M_6, M_7, M_8, M_9\}$; 7 in the Queens $\{Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, Q_7\}$, and 3 in the Staten Island $\{S_1, S_2, S_3\}$; 4 places do not belong to any borough $\{CW, T_1, X_2, X_4\}$. Figure 3 show the incident area in New York, and the ambulance dispatch area with amount of ambulance assigned in March and April 2020.

There are 10 codes indicating the incident dispositions seeing in Figure 4 (a). We can see the two cases that have the most percentage: '82 Transporting Patient' and '93 Refused Medical Aid'. Average total time of both two cases reach their highest level in March and April 2020 for the last five years.

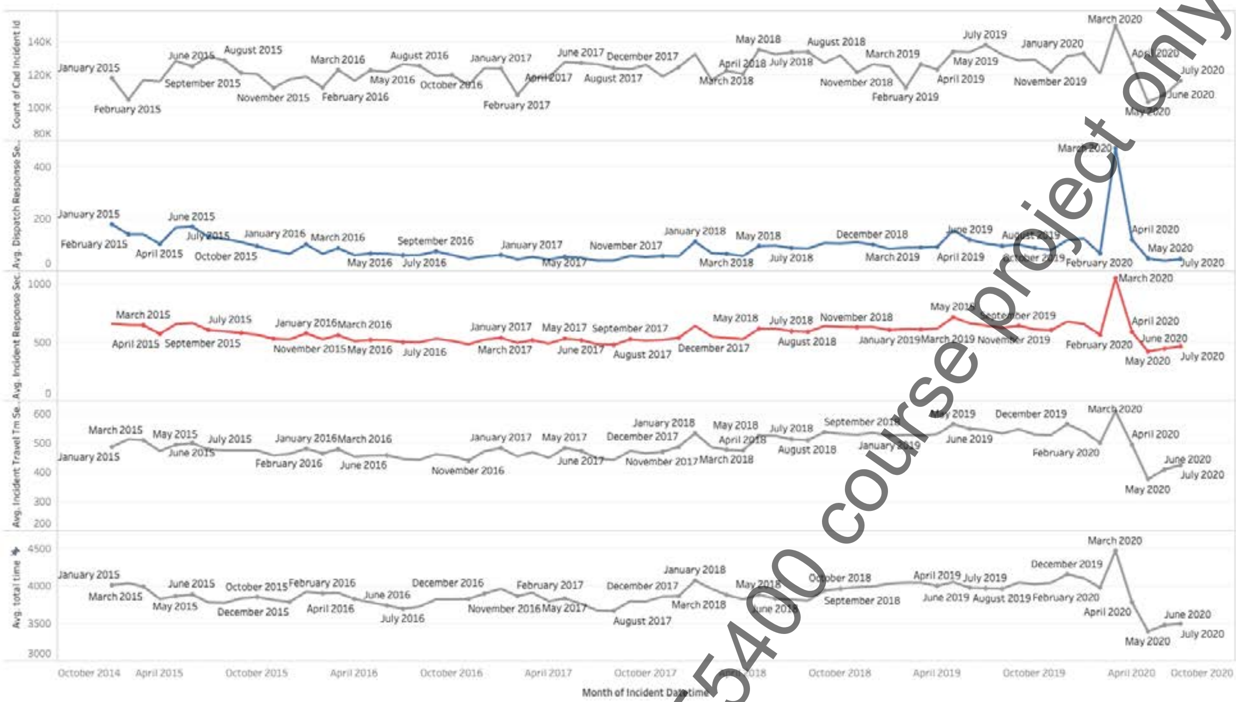


Figure 2: Incident Datatime Trend.

5. Proposed Model

5.1. Parameter and Variable Setup

Sets

I set of boroughs

T set of timestamps

K set of dispatch areas

d_{it} demand of borough i at time t.

a_k maximal resource at dispatch area k is able to serve.

b_{ki} trip availability from dispatch area k to borough i.

v_{ki} travel time serving borough i from dispatch k.

α weight for number of resources.

μ weight for maximal number of resources assigned from each dispatch area.

λ weight for travel time assigned to different boroughs.

ω weight for the balance of travel time between different boroughs.

Variables

x_{kt} number of resources assigned from dispatch area k at time t.

w_{kit} number of resources serving borough i from dispatch k at time t.

c_{it}^{+} dummy variables for linearzing absolute values in objective function.

c_{it}^{-} dummy variables for linearzing absolute values in objective function.

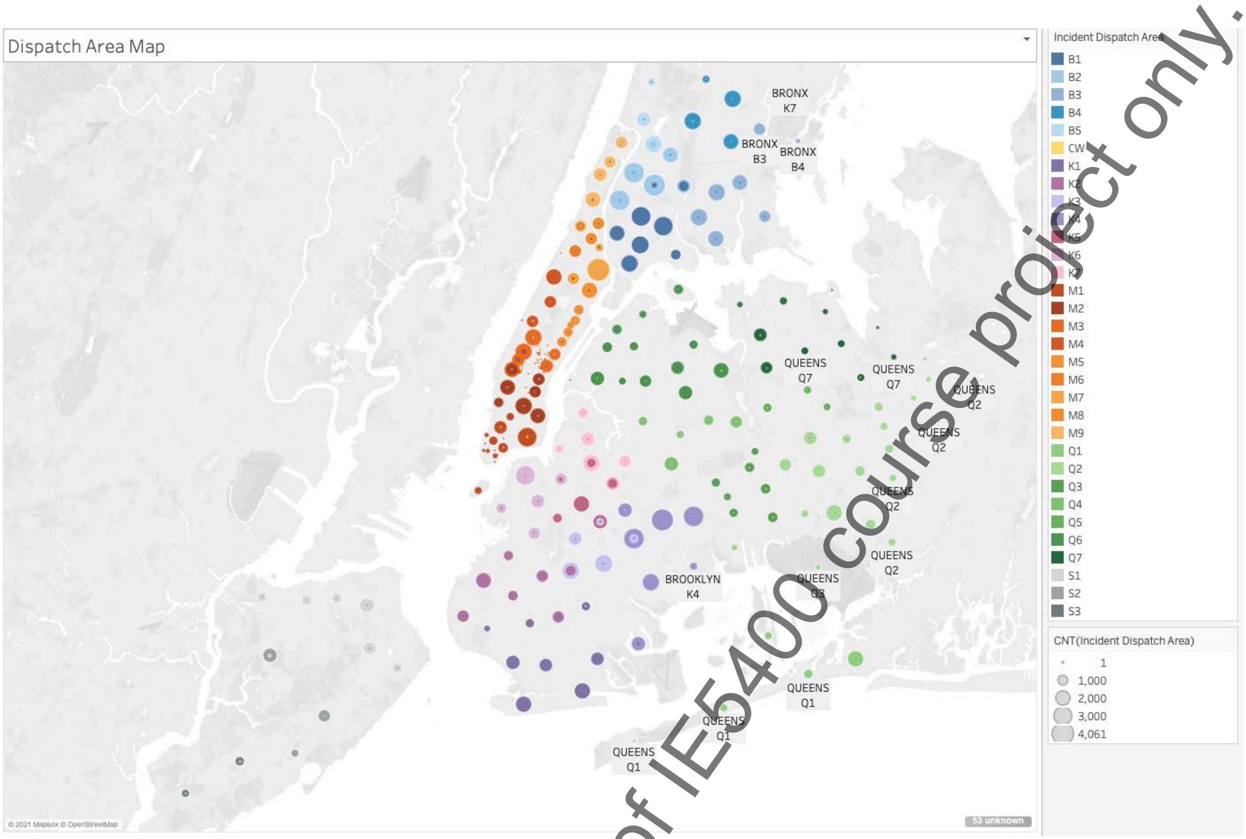
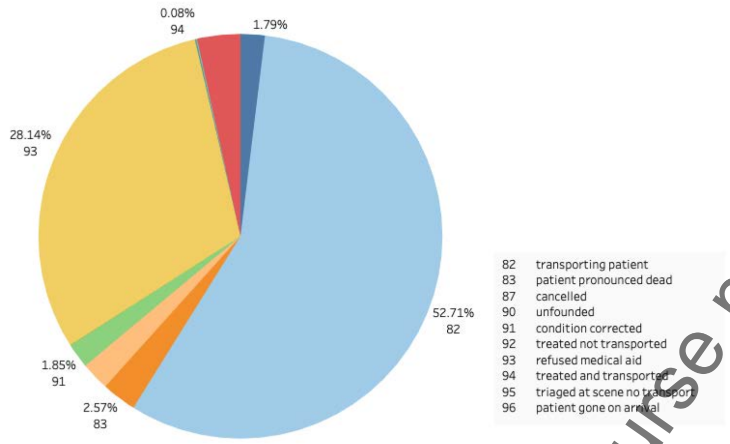


Figure 3: Dispatch Area and Amount in New York.

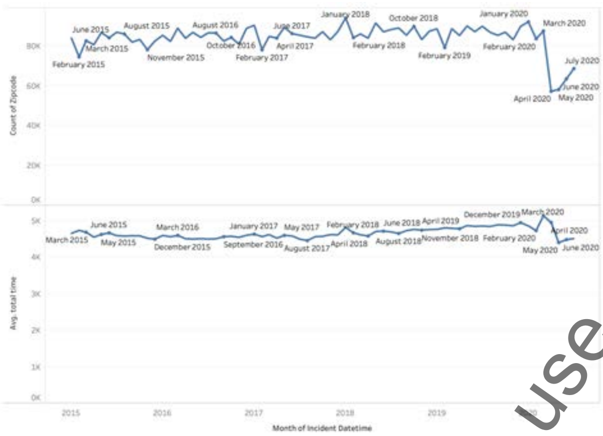
5.2. EMS Allocation-Reallocation Model

$$
\begin{aligned}
\text{(Model1)} \quad Min \quad & \alpha \sum_{k\in K, t\in T} x_{kt} + \lambda \sum_{k\in K, i\in I, t\in T} v_{ki} w_{kit} \\
& + \omega \sum_{i\in I, j\in I, t\in T, i\neq j} |v_{ki} w_{kit} - v_{kj} w_{kjt}| && (1) \\
s.t. \quad & x_{kt} = \sum_{i\in I} w_{kit} \quad \forall i \in I, k \in K, t \in T && (2) \\
& \sum_{i\in I} w_{kit} \leq \mu a_k \quad \forall k \in K, t \in T && (3) \\
& w_{kit} \leq \mu b_{ki} a_k \quad \forall k \in K, t \in T && (4) \\
& d_{it} = \sum_{k\in Q_i} w_{kit} \quad \forall k \in Q_i, e \in E_i, i \in I, t \in T
\end{aligned}
$$

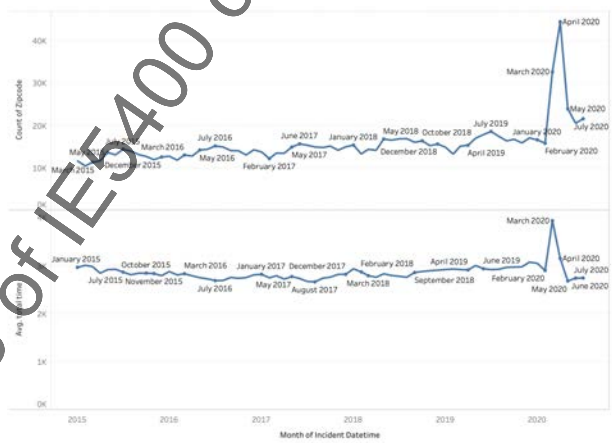
(5)



(a) Incident Dispositions.



(b) 82 Transporting Patient.



(c) 93 Refused Medical Aid.

Figure 4: Incident Dispositions in March and April 2020.

The objective function (1) is to minimize the number of resource assigned, minimize travel time, and the balance between the gaps on travel time over boroughs.

Constraint (2): the total number of resources that servers other boroughs form dispatch area k equals to the number of resources assigned to boroughs i.

Constraint (3): represents the maximal assignment availability from dispatch area k based on list a_k.

Constraint (4): represents the assignment availability from dispatch area k to borough i based on matrix b_{ki}.

Constraint (5): the number of demand being satisfied at borough i equals to the number of resources that serves region i from dispatch areas k.

The Model 1 is non-linear due to the last two term in the objective function. We can linearize it by replacing it with the term $\omega \sum_{i\in I, j\in I, t\in T, i\neq j} |v_{ki}w_{kit} - v_{kj}w_{kjt}|$ and we can linearize it as the following model.