MobiAmbulance: Optimal Scheduling of Emergency Vehicles in Catastrophic Situations

Li Yan*, Shohaib Mahmud*, Haiying Shen*, Natasha Zhang Foutz†, Donald E. Brown‡,
Wie Yusuf§, Derek Loftis¶, Lucas Lyons‖, Jonathan L. Goodall‡ and Joshua Anton**
*Department of Computer Science, University of Virginia, USA
†McIntire School of Commerce, University of Virginia, USA
‡Department of Engineering Systems and Environment, University of Virginia, USA
§School of Public Service, Old Dominion University, USA
¶Center for Coastal Resources Management, College of William & Mary, USA
‖Charlottesville Fire Department, USA
**X-Mode Social, Inc, USA
Email: {ly4ss, sm6re, hs6ms, ynf8a, deb}@virginia.edu, jyusuf@odu.edu, jdloftis@vims.edu,
lyonsl@charlottesville.org, goodall@virginia.edu, josh@xmodesocial.com

***Abstract*—With recent experience in multiple large-scale disasters, it has been widely confirmed that the severity of a disaster is greatly dependent on the effectiveness of ambulance dispatching during disaster phase. However, previous base station (i.e., temporary or permanent hospital) based ambulance redeployment methods and dynamic ambulance scheduling methods cannot handle the ambulance dispatching problem in catastrophic situations. In this paper, we present *MobiAmbulance*: a human *Mobi*lity based *Ambulance* dispatching system that aims to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delay of the fulfilled requests. We studied a state-scale human mobility dataset and found that the change of vehicle flow rate can be utilized to determine the connection status between road segments, and the distribution of people in catastrophic situations is drastically different from that in normal situations. Then, we develop a method to determine the road network connection status and the set of road segments that can still be driven through by ambulances after disaster. Based on the updated road network graph, we develop an ambulance dispatching method based on weighted driving route to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delays of the fulfilled requests. Our trace-driven experiments demonstrate the superior performance of *MobiAmbulance* over other comparison methods.**

I. Introduction

Emergency Medical Services (EMS) that rely on the dispatching of emergency vehicles to pick up patients and transport them to hospitals are of pivotal importance, especially during and after disasters. In this paper, we simply use ambulances to represent emergency vehicles used for transporting patients to hospitals [1] With experience in disasters such as Hurricane Florence (September 12-15, 2018) and Hurricane Michael (October 7-16, 2018), it has been widely recognized that the severity of a disaster is greatly dependent on the effectiveness of EMS during disaster phase. Although a lot of casualties are caused by the disaster itself, many casualties are due to lack of timely medical aid during the "golden rescue hour" after the disaster [1], [2]. Therefore, the timely and effective scheduling method of ambulances under catastrophic situations is necessary. Catastrophic situations and disasters are interchangeable terms in this paper.

Many methods that determine the deployment locations of ambulance base stations and optimize the stand-by base station for each ambulance at different times have been proposed [3]–[10]. A base station here means a temporary location where ambulances can stand by, and it is usually a temporary or permanent hospital. Generally, the works focus on utilizing various models that consider multiple factors (e.g., traffic, ambulance availability, patient appearance) to determine the deployment locations of base stations or the base station an ambulance should stand by at a specific time. However, these methods are not applicable for catastrophic situations, in which the factors exhibit different phenomenon. For example, during a disaster, there may be a large number of patient pick-up requests over a short period of time and in more condensed areas than in normal situations, and the original deployment of base stations may not be suitable for standing by anymore.

In the meantime, several dynamic ambulance scheduling methods also have been proposed [11]–[13]. These methods pre-schedule the driving route of each ambulance at the start of a day based on historical appearance of patient pick-up requests. However, large-scale disaster usually cause some road segments in the road network to be broken (i.e., inaccessible or congested). These methods cannot update the driving route of the ambulance according to the real-time road network connection status and the changed distribution of patient pick-up requests after disaster. Therefore, it is a challenge to obtain the real-time road network connection status and distribution of patient pick-up requests after a disaster, and optimize the driving route of an ambulance to pick up an patient to be the one that has the minimum driving delay based on the obtained information.

To handle the challenge, we propose *MobiAmbulance*, a human *Mobi*lity based *Ambulance* dispatching system that aims to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delay of the fulfilled requests. During catastrophic situations, real-time road connection status may not be updated fast enough due to interrupted traffic measurement. Human GPS positions can help predict potential ambulance requests and estimate the road network connection status. First, we analyzed a human mobility dataset that records the movements of most of the people in Charlotte, North Carolina before, during, and after the Hurricane Florence (September 12-15, 2018). We have the following observations from our dataset analysis:

(1) We found that the effect caused by the disaster can be described by the change of the road segment's vehicle flow rate (i.e., average number of passing vehicles per unit time). Thus, we can utilize the change of vehicle flow rate to determine the connection status of road segments and obtain the set of road segments that can still be driven through by a vehicle. The road connection status will serve as the foundation of the ambulance dispatching.
(2) We also found that the disaster greatly affected the mobility of people. In addition to doing routing that considers the road network connection status, the ambulance dispatching method must also adapt to the changed distribution of people.

The observations serve as the foundation for the design of *MobiAmbulance*. Accordingly, we first develop a method to determine the road network connection status and the set of road segments that can still be driven through by vehicles after disaster (called updated road network graph) by comparing the vehicle flow rates of the road segments before and after diaster. Then, based on the updated road network graph, we develop an ambulance dispatching method based on weighted driving routes to guide the ambulances. The ambulance dispatching method determines the driving route for each ambulance to maximize the ambulances' weighting function, which is defined in a way that increases the number of potential patient pick-up requests that all ambulances can possibly fulfill, and reduces the total driving delay that the ambulances need to drive to their newly assigned road segments. In summary, our contributions include:

1. Our analysis on a state-scale human mobility dataset confirms the effect of disaster on vehicle flow rate and human movement, and aids in the design of *MobiAmbulance*.
2. We propose *MobiAmbulance*, a human *Mobi*lity based *Ambulance* dispatching system that aims to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delay of the fulfilled requests.
3. We have conducted extensive trace-driven experiments to show the effectiveness of *MobiAmbulance* in terms of the number of fulfilled patient pick-up requests per unit time, and driving delay of fulfilling the requests.

To our knowledge, this paper is the first work for ambulance dispatching under dynamic catastrophic situations. The remainder of the paper is organized as follows. Section II presents literature review. Section III presents our dataset analysis results. Section IV presents the detailed design of *MobiAmbulance*. Section V presents performance evaluations. Section VII concludes the paper with remarks on future work.

II. Related Work

Base station based ambulance redeployment. Many methods [3]–[10] that determine base stations for ambulances based on the dynamic change of patient pick-up requests have been proposed. Yue *et al.* [3] proposed considering the allocation and dynamic redeployment of ambulances by using simulation-based approach to find the ambulance deployment strategy that has the minimum driving delay. Schmid *et al.* [4] formulated a mixed integer programming model to maximize the total patient pick-up request coverage of ambulance deployment and also minimize the ambulances' driving delays. Saisubramanian *et al.* [5] generated an ambulance deployment strategy that minimizes the ambulances' driving delay with a bounded risk (i.e., percentage of incidents that can have driving delays higher than the objective) using a linear optimization model. Gendreau *et al.* [6] proposed the dynamic double standard model (DDSM) that solves the ambulance redeployment problem based on the objective of the double standard model (DSM). Degel *et al.* [7] developed a multi-period model for EMS location and relocation planning at a tactical level in which the temporal and spatial variations in demand and driving delays are captured. Van *et al.* [8] proposed an integer programming based ambulance deployment model that focuses on minimizing the average driving delay of the ambulances to the patients' pick-up request locations. Snyder *et al.* [9] proposed several reliability models to find the optimal location of ambulance base stations so as to minimize regular driving cost of the ambulances when some of the base stations do not have available ambulances. Daskin *et al.* [10] formulated the maximum expected covering location model (MEXCLM) and a heuristic solution to maximally cover the potential patient pick-up requests with the minimum driving delays. However, these methods are not applicable for catastrophic situations, in which the factors may be relatively more abnormal and the original deployment of base stations may not be suitable for standing by.

Dynamic scheduling of ambulances. Many methods [11]–[13] that focus on scheduling the real-time cruising route of ambulances to further reduce the ambulances' driving delay to the target patients' pick-up positions have been proposed. Schmid *et al.* [11] proposed a stochastic dynamic model that combines dynamic ambulance relocation from an ambulance's real-time position and updated driving route to minimize the ambulance's driving delay. Maxwell *et al.* [12] proposed an approximate dynamic programming (ADP) model that schedules the driving routes of idle ambulances such that the real-time coverage of potential patient pick-up requests is maximized. Jagtenberg *et al.* [13] developed a heuristic algorithm for real-time ambulance redeployment based on MEXCLP. The dispatching policy for idle ambulances is based on maximizing the marginal expected coverage of the potential patient pick-up requests. However, these methods cannot update the driving route of the ambulance according to the real-time road network connection status and the changed distribution of patient pick-up requests after disaster.

III. Metropolitan-Scale Dataset Measurement

A. Dataset Description and Definitions

Our datasets, provided by X-Mode [14], record the human mobility of people in Charlotte city of North Carolina State 15 days prior and after the Hurricane Florence (from September 1 to September 24, 2018).

The datasets include:

- **GPS Data.** GPS data is collected from individual's cell phone's GPS sensor at some irregular time interval (from 0.5 hours to 2 hours). The data contains time-stamp, latitude, longitude, altitude, and speed of each user during the sampling period from September 1 to September 24, 2018. The associated time stamp and unique id of each individual allows us to track each user anonymously. The location data covers all over the Charlotte city. The data was captured from September 1 to September 24 encompassing the development of the hurricane Florence (September 1 to 11), landfall-when the center of the eye of hurricane reaches land (September 14) and weakening (September 16 to 19) of the hurricane Florence.
- **Road map data.** The road network map of Charlotte city is obtained from OpenStreetMap [15]. We have used a bounding box with coordinate ($lat = 35.6022, lon = -79.0735$) as south-west corner, and coordinate ($lat = 36.0070, lon = -78.2592$) as north-east corner. We have used the data from National Weather Service to crop the affected area.

For data management, we utilized a 34TB Hadoop Distributed File System (HDFS) [16] on a cluster consisting of 10 nodes, each of which is equipped with 16 cores and 64 GB RAM. For data processing, we used Apache Spark [17], which is a fast in-memory cluster computing system running on Hadoop. We explore spatio-temporal correlation and divergence between the mobility (in terms of trips) obtained by cellphones. Such correlation and divergence serve as the empirical guidelines for our later inference.

We represent the road network of Charlotte city with a directed graph $G = (E, V)$, in which vertices V represent landmarks (i.e., intersections or turning points), and edges E represent road segments [18]–[21]. Based on the road network, we introduce the following definition:

Definition 1. Human Trajectory. *A human's trajectory is a sequence of N^e time-ordered landmarks, where each landmark is represented by a latitude and a longitude.*

Definition 2. Vehicle Flow Rate. *Vehicle flow rate of a road segment is defined as the average number of vehicles driving through the road segment per unit time [18].*

B. Dataset Analysis

1) ***Observation 1****: Vehicle Flow Rate Before and After Disaster:* Generally, the incidents caused by a disaster (e.g., wreckage, flooding, traffic accidents) will severely change the traffic status of many road segments. To ensure the service efficiency of ambulances under catastrophic situations, the ambulance dispatching system must reliably offer the driving route that let the ambulances drive through to their target patient pick-up locations without confronting obstacles. Vehicle flow rate, which measures the number of vehicles passing through a road segment per unit time, is a direct metric showing whether the road segment is still suitable for an ambulance to drive through after disaster breakout. To illustrate the the road segment connection status and the vehicle flow rate of the road segment, we measured the vehicle flow rate of a flooded road segment per hour on August 25, 2018 (i.e., before disaster) and September 20, 2018 (i.e., after disaster).

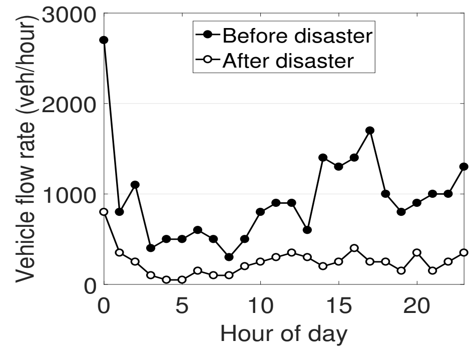

Fig. 1: Vehicle flow rate of a road segment before and after disaster.

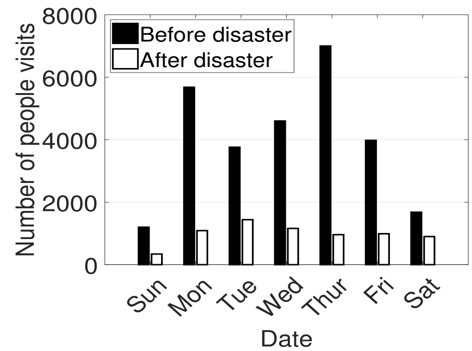

Fig. 2: Distribution of people in a CBD area before and after disaster.

Figure 1 shows the measured results. We can see that the gap between the vehicle flow rates of this road segment before and after the disaster can be as high as 1,500 (at 00:00) and as low as 200 (at 08:00). The result shows that the effect caused by the flooding can be described with the change of the road segment's vehicle flow rate. Thus, we can utilize the change of vehicle flow rate to determine the connection status of road segments and obtain the set of road segments that can still be driven through by an ambulance. In Section IV-B, we will elaborate the details of our method on estimating the road network connection status.

2) ***Observation 2****: Distribution of People Before and After Disaster:* Due to the effect of disaster, the distribution of people may also significantly change, which invalidate the ambulance dispatching methods that are designed for normal situations. We consider that a person has visited an area if his/her location stays in the area for more than 5 minutes. To confirm this conjecture, we measured the numbers of people visits to Charlotte's Central Business District (CBD) in the 7 days of a week before the disaster (August 19–August 25) and in the 7 days of a week after the disaster (September 18–September 24), respectively .

The results are illustrated in Figure 2. The numbers of people visits during the days before the impact of disaster are much higher than those during the days after the impact of disaster. The gap in the numbers of people visits between "Before disaster" and "After disaster" can be as high as 5,000. This shows that the disaster also greatly affected the mobility of people. In addition to do routing considering the road network connection status, the ambulance dispatching method must also adapt to the changed distribution of people. In Section IV-C, we will elaborate the details of our ambulance dispatching method based on weighted driving route. Based on these observations, we will find a solution in Section IV for the following problem.

Problem statement: Given a road network G affected by a diaster, and human mobility data, how to determine the road network connection status after disaster, and utilize the obtained road network connection status to dispatch the ambulances to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delay of the fulfilled requests?

We assume that the capacity of an ambulance is c, which can be set by the ambulance dispatching center according to actual equipment details (e.g., c =1, 5, etc.). There will be multiple rounds of dispatches within 24 hours per ambulance. Once the patients are transported to a certain hospital, it will be dispatched to serve the next patient pick-up request.

IV. System Design Details

A. Framework of MobiAmbulance

MobiAmbulance consists of the following stages:

1. **Human mobility information derivation** (Section III-A). First, we apply the *Data Cleaning* (e.g., filtering out positions out of the actual range of our interested city, redundant positions). Then, based on the cleaned data, we derive the *Trajectories in Landmarks* of humans on the *Roadmap with Landmarks and Road Segments*.
2. **Road network connection status estimation** (Section IV-B). Based on the output *Trajectories in Landmarks* from the first stage, we estimate the *Road Network Connection Status* to obtain the set of available road segments that the ambulances can still drive through after disaster.
3. **Ambulance Dispatching based on Weighted Driving Route** (Section IV-C). Based on the obtained set of available road segments, we utilize the *Ambulance Dispatching based on Weighted Driving Route* to decide the road segment each ambulance should drive to maximize the total number of potential patient pick-up requests that can be served by the ambulances, and minimize the ambulances' driving delay in fulfilling these requests.

B. Road Network Connection Status Estimation

After disaster, the road network is usually broken by certain issues such as flooding, wreckage, etc. Therefore, it is challenging to know the real-time road network connection status that the ambulances can drive through. This section presents our road network connection status estimation method that resolves the problem: *how to obtain the real-time road network connection status after disaster?* In the following sections, we first specify how *MobiAmbulance* determines the real-time road network connection status (Section IV-B1). Second, we will specify how *MobiAmbulance* determines the set of the remaining road segments that the ambulances can drive through (called updated road network graph) to fulfill patients' pick-up requests (Section IV-B2).

1) Determining The Real-time Road Network Connection Status: Recall that the conclusion from Observation 1 (Section III-B1) indicates that the road segments with significant change of vehicle flow rate. Note that according to Definition 2, vehicle flow rate of a road segment is defined as the average number of vehicles driving through the road segment per unit time. After disaster, the vehicle flow rate may significantly decrease due to broken road segment, traffic jam, etc. Through comparing the vehicle flow rate of a certain road segment before and after the disaster, we can generally determine whether an ambulance can drive through the road segment. If we do this comparison for each road segment of the road network, we can determine the real-time road network connection status. Specifically, we define the measurement period of the vehicle flow rate of a road segment as the time interval of measuring the number of vehicles that pass through the road segment. Recall that we partition a day into several time slots. We let a measurement period equal to a time slot without losing generality.

After every measurement period (e.g., 15 minutes), we obtain the vehicle flow rate of each road segment (called the current vehicle flow rate) and compare the result with the vehicle flow rate during the same measurement period in the most recent day without disaster (called the normal vehicle flow rate). Due to various factors (e.g., flooding, debris, traffic jam), the vehicle flow rate of some road segments will greatly decrease. If the current vehicle flow rate is much lower than the normal vehicle flow rate and the difference is larger than a threshold, it means that the traffic of this road segment is significantly affected and becomes not suitable for an ambulance to drive through. The comparison method can be represented as:

$$\Delta f = f_n - f_c < \gamma, \tag{1}$$

where f_n represents the normal vehicle flow rate, f_c represents the current vehicle flow rate, Δf is the measured difference between f_n and f_c, and γ is the vehicle flow rate threshold to determine if the vehicle flow rate change is large. The value setting of γ controls the tolerance degree of the vehicle flow rate change. If the value of γ is relatively higher, the comparison can tolerate a certain vehicle flow rate increase and there will be more route options, but the obtained suitable road segments may result in longer driving delay for the ambulances to drive through, and vice versa. Thus, the value of γ should be set according to ambulances' requirements on driving delay.

2) Determining the Updated Road Network Graph: The broken connections between road segments in a disaster need to be removed from the original road network graph to create the updated road network graph. By applying the comparison

Algorithm 1: Determining the updated road network graph.

Input : obtained set of road segments $\widetilde{G}$
Output: updated road network graph G'

1 utilize the beginning of a randomly selected road segment $e_i \in \widetilde{G}$ as start position;
2 initialize the minimum driving delay $t_{\text{min}} = \text{Inf}$;
3 initialize the Euler path with the minimum driving delay $r_{\text{min}} = \text{None}$;
4 **while** $e_j \in \widetilde{G}$ has not been selected as start position **do**
5   utilize the beginning of e_j as start position;
6   apply the LKH algorithm on $\widetilde{G}$ to obtain an Euler path r_j;
7   calculate the driving delay of r_j: t_j;
8   **if** $t_j < t_{min}$ **then**
9   assign t_j to t_{min};
10   assign r_j to r_{min};
11 **return** the set of road segments covered by r_{min} as G'

Equation (1) on each road segment, we can obtain the set of all the road segments that the ambulances can drive through. Since the obtained set of road segments only covers partial road network, so we denote it as $\widetilde{G} = (\widetilde{E}, \widetilde{V})$, where $\widetilde{E}$ and $\widetilde{V}$ represent the partial road segments and partial landmarks included in $\widetilde{G}$, respectively. However, these road segments may not be connected to each other due to certain broken road segments between them. To connect these road segments, we need to add back several road segments that may have been judged as "broken" by Equation (1). The additional road segments must be the ones that result in the shortest driving delay for ambulances. Therefore, in this section, we explain how *MobiAmbulance* finds the updated road network graph that covers all the road segments in $\widetilde{G}$ and a driving path that passes through all the road segments in $\widetilde{G}$ only once and has the minimum total driving delay for ambulances.

Specifically, to find the updated road network graph that covers all the road segments in $\widetilde{G}$ and a driving path that passes through all the road segments in $\widetilde{G}$ only once so there is always a path given a source-destination pair, we actually need to find an Euler path [22] through $\widetilde{G}$, which covers all the road segments of $\widetilde{G}$ but only once for each segment, by applying the LKH (Lin-Kernighan heuristic) algorithm [22]. There may be multiple Euler paths that fulfill this requirement, and we need to select the one that results in the shortest driving delay. Starting from the beginning of a randomly selected road segment in $\widetilde{G}$, we utilize the LKH algorithm to obtain the shortest Euler path that covers $\widetilde{G}$. We repeat this procedure until all the road segments of $\widetilde{G}$ have been selected as the beginning road segment once. Finally, we extract the Euler path that results in the shortest total driving delay based on the road segments' current traffic as the updated road network graph.

The whole process can be summarized by Algorithm 1. At Line 1, the algorithm initializes the start position to obtain the initial Euler path. From Line 1 to Line 3, we initialize the start position of the initial temporary Euler path, the minimum driving delay over all the attempted Euler paths t_{min}, and the corresponding Euler path r_{min}. From Line 4 to Line 10, we iterate all the road segments included in $\widetilde{G}$ as the start position to generate a temporary Euler path, compare the Euler path's driving delay t_j with the recorded minimum driving delay t_{min}, and update the Euler path with the minimum driving delay r_{min}.

C. Ambulance Dispatching based on Weighted Driving Route

From Observation 2 (Section III-B2), we know that the distribution of people has significant change before and after disaster, and then we can infer that the distribution of patient pick-up requests may have significant change before and after disaster correspondingly. The previous ambulance dispatching methods handling the normal situations [11]–[13] cannot ensure that the patients' pick-up requests can be fulfilled in time under catastrophic situations. Thus, to solve the problem, we need a method to dispatch the ambulances to pick up the patients with the minimum driving delay and meanwhile maximize the number of picked up patients in a global manner according to the appearance of pick-up requests.

Therefore, we propose an ambulance dispatching system that determines the action of each ambulance based on the estimated weights of candidate driving routes. The action of an ambulance will be driving to which road segment to serve patients' pick-up requests. The patients' pick-up requests are reported to the ambulance dispatching center for the global routing of the ambulances. The guidance process of the model is that given a current road network connection status and distribution of patients' pick-up requests, it outputs an action for each ambulance. The actions of all the ambulances enable them to drive the shortest route with the minimum driving delay and fulfill the maximum number of potential patient pick-up requests.

We define the *weight* caused by all the ambulances driving through a route as the weighted sum of the total number of potential patient pick-up requests that the ambulances can serve, and the sum of the ambulances' driving delay to serve the requests. We define the driving delay of an ambulance as the ambulance's traveling time from its current position to the end of its destination road segment determined by the ambulance dispatching system. The ambulance dispatching system updates the driving route for each of the ambulances to fulfill the patients' pick-up request with the shortest overall driving delay periodically (e.g., every 15 minutes). The guided ambulances will serve the pick-up requests appearing on the road segments in their driving routes. In the following, we introduce the details of the ambulances' action, the number of encountering patient pick-up requests and weighting function of driving route for the ambulance dispatching system.

1) Ambulances' Action: The action output by the ambulance dispatching system provides a driving decision for each ambulance. Specifically, the k^{th} ambulance's driving decision (denoted by x_k) means the destination road segment and it can be any road segment in the road network ($e_j \in E$). If

$x_k = e_j \in E$, we use an existing shortest distance routing method (e.g., Dijkstra algorithm [18]) to determine the k^{th} ambulance's driving route from its current position to the end of its destination road segment e_j. Thus, the action can be represented as:

$$a = (x_k \mid \forall\, m_k \in \mathcal{A}), \tag{2}$$

where a represents the action, which includes the driving decision of all the ambulances x_k represents the driving decision of the k^{th} ambulance, and m_k represents the k^{th} ambulance.

2) Number of Encountering Patient Pick-up Requests: Since the ambulance dispatching system needs to determine which road segment each ambulance needs to drive to in order to maximize the total number of fulfilled patients' pick-up requests of all the ambulances, we consider the estimated number of potential patients that each ambulance may encounter on route if it is chosen to drive to the end of each road segment in the road network. We take the k^{th} ambulance (denoted by m_k) as an example. For each road segment in the road network ($e_j \in E$), we use the Dijkstra algorithm [18] to find the shortest distance trajectory from the ambulance's current position (denoted by p_k) to the end of road segment e_j. The trajectory consists of a sequence of road segments from p_k to the end of e_j, and is denoted as $\Phi_{kj} = \{p_k, \ldots, e_j\}$. By this way, we calculate the number of potential patient pick-up requests in each vehicle's trajectory.

Specifically, based on the traffic of the road segments in the trajectory, we deduce the general arrival timestamp t_i to each of the road segments e_i in the trajectory. That is, the ambulance's driving delay to pick up the patients on e_i is t_i. Suppose the number of potential patient pick-up requests appearing on e_i is n_{e_i}. Finally, we use the sum of the number of potential patient pick-up requests of all the road segments $\mathcal{N}_k = \sum_{e_i \in \Phi_{kj}} n_{e_i}$ as the k^{th} ambulance's estimated number of potential patient pick-up requests if it drives to e_j. Thus, when dispatching the ambulances, we consider each ambulance's estimated numbers of potential patient pick-up requests on the ambulance's trajectory to each road segment. It can be represented as follows:

$$s = (\mathcal{N}_k \mid \forall\, m_k \in \mathcal{A}) \tag{3}$$

where s represents the number of encountering patient pick-up requests of all the ambulances. $\mathcal{N}_k = \{\sum_{e_i \in \Phi_{kj}} n_{e_i} | e_j \in E\}$ is the estimated potential number of patient pick-up requests that the k^{th} ambulance can encounter if it drives to e_j. $e_j \in E$ represents a candidate road segment belonging to the total set of road segments E in the road network. m_k represents the k^{th} ambulance. $\mathcal{A}$ represents the set of all the ambulances. When determining the driving route, we need to avoid duplicate coverage of patient pick-up requests on the same road segment. By duplicate, we mean that if the total capacity of dispatched ambulances can already cover all the requests on a certain road segment, there is no need to further dispatch more ambulances to cover the road segment. This is because that driving duplicate road segments will waste the ambulances' ability in serving requests.

For example, suppose an ambulance with a capacity of $c = 5$ has been dispatched to drive the route e_1(1 requests) $\rightarrow$ e_2(2 requests) $\rightarrow$ e_3(1 requests). Since $c = 5 > 4$, all the 4 requests on e_1, e_2 and e_3 can be covered by the ambulance. Therefore, the other ambulances should not drive through these road segments again. Therefore, when determining the driving route of an ambulance, we exclude the road segments whose patient pick-up requests have already been covered by the capacity of dispatched ambulances and let the ambulance only drive the road segments whose requests have not been fully covered.

3) Weighting Function: Since we expect the patient pick-up requests can be fulfilled by the ambulances' service as early as possible, the ambulances' driving delay should be minimized. Therefore, the main goal of the ambulance dispatching system is to find the driving routes of the ambulances to maximize the total number of potential patient pick-up requests that can be served by the ambulances, and meanwhile minimize the sum of all the ambulances' driving delays. The weighting function resulted by the x_k of all the ambulances is defined as the weighted sum of the total number of potential patient pick-up requests that the ambulances can serve, and the sum of the ambulances' driving delays, which can be formulated as:

$$r(s_t, a_t) = \alpha N^q - \beta T^d \tag{4}$$

where s_t is the number of encountering patient pick-up requests of all the ambulances at current time t, a_t is the driving actions of all the ambulances at current time, N^q and T^d are the two metrics affecting the weighting function. N^q denotes the total number of patient pick-up requests metric contributed by all the ambulances. The more requests all the ambulances will encounter on route, the higher weight the ambulances will contribute. T^d denotes the driving delay metric of the ambulances. The shorter driving delay all the ambulances will have, the higher weight the ambulances will contribute. α and β are constants that control the respective influence of the metrics and can be adjusted manually. Below, we explain how we calculate N^q and T^d.

The total number of fulfilled patient pick-up requests metric is the total number of patient pick-up requests that all the ambulances will encounter by driving to different destination road segments. When making the driving decision for each ambulance, as aforementioned, we need to avoid duplicate coverage of patient pick-up requests on the same road segment. This is because that driving duplicate road segments will waste the ambulances' ability in serving patient pick-up requests. Therefore, we need to globally count the total number of fulfilled patient pick-up requests contributed by all the ambulances. Specifically, suppose $E' = \bigcap_{m_k \in \mathcal{A}} \Phi_{kj}$ is the intersection set of the road segments that all the ambulances can cover by making their respective driving decisions. We calculate the sum of the numbers of patient pick-up requests of all the road segments in E' as the total number of fulfilled patient pick-up requests metric. This metric is calculated as:

$$N^q = \sum_{e_i \in E'} n_{e_i}, \tag{5}$$

where E' is determined by the driving decision of each ambulance (x_k), which is from the action a_t, e_i is an element road segment covered in E', and $\sum_{e_j \in E'} n_{e_j}$, which is from Section IV-C2, is the estimated potential number of fulfilled patient pick-up requests of the ambulances that are chosen to drive to e_j. The more potential patient pick-up requests the ambulances will encounter, the higher weight the ambulances will contribute.

The ambulance driving delay metric is the sum of all the ambulances' driving delays $T^d = \sum_{m_k \in \mathcal{A}} t_k^d$, where t_k^d is the driving delay of the k^{th} ambulance. Specifically, the k^{th} ambulance's driving delay to e_j if it is chosen to drive to the end of the road segment e_j ($x_k = e_j \in E$), is calculated by $t_{kj} = \sum_{e_i \in \Phi_{kj}} \frac{l_{e_i}}{v_i}$, where l_{e_i} is the length of e_i, and v_i is the speed limit of e_i under current catastrophic condition.

Finally, by following a certain dispatching period (e.g., 15 minutes), we iterate all the candidate road segments for each ambulance and determine the driving routes to the target destination road segment of each ambulance that maximize the weighting function. That is, maximize $r(s_t, a_t)$.

V. Performance Evaluation

A. Comparison Methods

To evaluate *MobiAmbulance*'s performance (*MA* in short), we compare its performances of fulfilling patient pick-up requests and driving delay of fulfilling these requests with a representative base station based ambulance redeployment method [5] (*Deploy* in short), and a representative dynamic ambulance scheduling method [13] (*Schedule* in short).

Specifically, *Deploy* aims to minimize the ambulances' driving delay with a bounded risk (i.e., percentage of incidents that have driving delays higher than a certain objective), and deploys the ambulances to their respective optimal base station (i.e., hospital) to cover emerging patient pick-up requests. Once the ambulances are deployed to the optimal base station, the available ambulance nearest to the request will drive to the request position to serve the request. The drawback of this method is that the deployment of hospitals may not be suitable for standing by during catastrophic situations and may cause long driving delays.

Schedule considers the distribution of the positions of patient pick-up requests and aims to maximize the real-time coverage of the requests through dynamically scheduling the ambulances' driving routes. The drawback of this method is that it cannot update the driving route of the ambulance according to the real-time road network connection status and the changed distribution of patient pick-up requests after disaster. Driving the "broken" road segments may cause the ambulances to have high driving delays. For fair comparison, we suppose the default deployment of hospitals in the three methods follows the deployment of existing hospitals in Charlotte city.

B. Experiment Settings

We use the human mobility data described in Section III-A to simulate the appearance of patient pick-up requests. The value of γ affects the accuracy of determining whether a certain road segment is broken. To find the best value of γ, we vary the value within a certain range (e.g., [100 veh/hour, 500 veh/hour]) and test the performance. Then, we choose the value that results in the minimum driving delay to the patient pick-up requests as the final setting. We find $\gamma = 250$ veh/hour is the best value for the case of Charlotte city. We suppose that a patient pick-up request expires after 30 minutes since its appearance. We assume that the ratio of patients over the whole population in Charlotte city is 5%. That is, during each time, 5% of the population becomes patients and needs ambulance service. We also suppose the pick-up requests of these patients randomly appear within respective time slots. Based on the deployment of existing hospitals in Charlotte city, we used SUMO [23] to simulate the movement of 500 ambulances for 24 hours on Charlotte's road network. We converted OpenStreetMap road network of Charlotte city to a SUMO road network file.

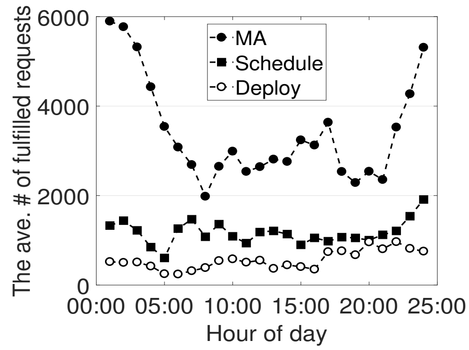

Fig. 3: The average number of fulfilled patient pick-up requests.

The metrics we measured are:

- *The average number of fulfilled patient pick-up requests*: For each ambulance, we measure the number of patient pick-up requests it fulfilled in each time slot throughout a day. Then, we measure the average number of fulfilled requests over all the ambulances in each time slot. The purpose of this metric is to compare the performance of different methods in covering the patient pick-up requests.
- *The average driving delay of the fulfilled requests*: For each fulfilled request, we measure the driving delay of the ambulance that fulfills this request. Then, we measure the average driving delay over all the fulfilled requests in each time slot. The purpose of this metric is to compare the performance of different methods in reducing the waiting time of the patients before being picked up.

C. Experimental Results

1) The average number of fulfilled patient pick-up requests: Figure 3 shows the average number of fulfilled patient pick-up requests over all the ambulances during each time slot throughout a day under different methods. We can see that the results of *MobiAmbulance* is always much higher than the other methods. The results follow: *MobiAmbulance>Schedule>Deploy*.

In *Deploy*, the ambulances are deployed to wait at their optimal base stations. Once a patient pick-up request appears, the available ambulance nearest to the request will drive to the request position. There is no specific method to dynamically

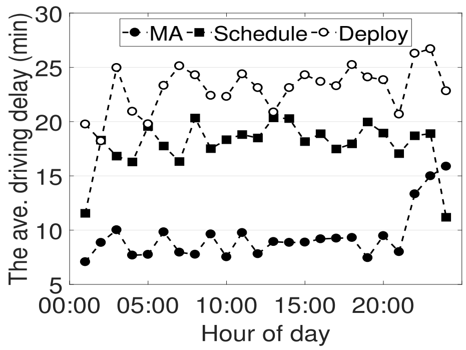

Fig. 4: The average driving delay of the fulfilled requests.

arrange the driving route to let the ambulances keep approaching the locations of the pick-up requests. The ambulances will only be dispatched to certain base stations (i.e. hospitals) to cover the requests appear nearby the hospital. Due to the effect of disaster, many potential patients were trapped in somewhere distant away from the coverage of the base stations. Therefore, the ambulances in *Deploy* missed the most patient pick-up requests.

In *Schedule*, the ambulances' driving routes are dynamically updated according to the appearance of potential patient pick-up requests. The ambulances are always dispatched to the driving routes that can cover the appearance locations of the requests. Therefore, the ambulances in *Schedule* fulfilled much more requests than those in *Deploy*. However, in determining the driving route, *Schedule* cannot offer guidance based on the actual road network connection status, but determine driving route based on the road network status in normal situations. This caused that some ambulances cannot reach their target potential patients within the 30 minute expiration limit, and still missed many requests.

In contrast, the updated road network graph generated by *MobiAmbulance* offers the current connection status of the road segments. When dispatching the ambulances, *MobiAmbulance* can avoid the ambulances from driving through the routes that will cause long driving delays. Also, the ambulance dispatching system can always guide the ambulances to cruise around the areas with possible appearance of potential patient pick-up requests. Therefore, the ambulances of *MobiAmbulance* can fulfill the most patient pick-up requests among the methods.

2) The average driving delay of the fulfilled requests: Figure 4 shows the average driving delays of the fulfilled patient pick-up requests during each time slot throughout a day under different methods. We can see that the results of *MobiAmbulance* are much shorter than the other methods during most time slots. The results follow: *MobiAmbulance*<*Schedule*<*Deploy*. These results are generally consistent with those demonstrated in Figure 3. In *Deploy*, the ambulances are deployed to wait at base stations before driving to the patient pick-up request locations. Once a patient pick-up request appears, the available ambulance nearest to the request will drive to the request position. As mentioned in Section V-C1, the hospitals are deployed based on normal patient pick-up request appearance. During catastrophic situations, the patient pick-up requests may not appear as in the normal situations, so the ambulances have to spend the longest time to transit from base stations to their target request locations.

Fig. 5: VIMS' street level model of maximum inundation extent along the Neuse River at Oriental, NC, 32 km (20 miles) east of New Bern along the Neuse River, where some of the highest precipitation measurements were recorded during 2018 Hurricane Florence.

VI. Future Design And Development Considerations

Recent advancements in the application of computational flood modeling technology have enabled hydrodynamic precision to hone in, calculate and predict the mass and movement of flood waters to accurately predict water velocities at the street-scale. Route guidance for emergency vehicles after a major storm event typically involves situational inference on behalf of the driver, and knowing which streets will be flooded. Larger localities such as Charlotte will likely have an environmental Geographic Information System (GIS) division that can model flooding affects. These considerations will be programmed into the dispatching and operational GIS maps in sophisticated systems. Cross-scale sub-grid inundation has seen a lot of improvements in the recent years which can now generate spatial inundation maps within 36 hours of a storm's arrival. For example, the Virginia Institute of Marine Science (VIMS) has developed a storm tide model and mapping tool called "Tidewatch". Figure 5 shows VIMS' street level model of maximum inundation extent along the Neuse River at Oriental, NC. Jointly considering the inferred flooding situation for scheduling of emergency vehicles is one of our future works. Morsy *et al.* [24] designed and prototyped a cloud-based system to support decision makers as they assess flood risk to transportation infrastructure during extreme weather events. The system automates access and preprocessing of forecast data, execution of a high-resolution 2D hydrodynamic model, and map-based visualization of model outputs.

MobiAmbulance has been designed with a focus to fulfill end users' (i.e., patients') pick up demand. Though experimental result of *MobiAmbulance* shows great technical feasibility, its effectiveness as a disaster response/recovery solution depends on how it will be utilized the users. Future design and development activities may emphasize a participatory design approach to incorporate users, i.e., emergency managers, emergency first responders (paramedics/EMTs), in the co-design process to make the *MobiAmbulance* more relevant and applicable to practice. Future work will enhance the utility of *MobiAmbulance* to end users by shifting the focus of next phase design and development activities to the settings where the technology will be used and embedding these activities in the "lived work" of users, thus taking into account users' needs, opinions, practices and habits [25] and social network analysis [26]. Leveraging the "lived experience" of the users and incorporating the intricacies of their decision making process will ensure greater acceptance of *MobiAmbulance* and increased likelihood of its utilization.

Another scope for improvement is utilizing GPS data associated with emergency vehicle/ambulance travel during and following a disaster. Many ambulances are equipped with GPS-based automatic vehicle locating (AVL) systems that provide real-time information about their movement. GPS data from these vehicles can be used to study their actual route selection [27]. This important user input about ambulance routing decisions "in practice" during disasters can further inform *MobiAmbulance* design and development.

VII. Conclusion

Ambulance dispatching during disaster phase is of great importance. Previous ambulance dispatching methods cannot effectively handle catastrophic situation. Our proposed *MobiAmbulance* is the first human mobility based ambulance dispatching system that utilizes vehicle flow rate comparison and ambulance dispatching based on weighted driving route to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delay of the fulfilled requests. Our analytical results on a state-scale human mobility dataset provide foundation for the design of *MobiAmbulance*. We develop a method that utilizes the comparison of vehicle flow rate to determine the road network connection status and the set of road segments that can still be driven through by ambulances after disaster (called updated road network graph). Based on the obtained updated road network graph, we develop an ambulance dispatching method based on weighted driving routes to maximize the total number of fulfilled patient pick-up requests, and minimize the driving delays to fulfill the pick-up requests. We conducted trace-driven experiments on SUMO to verify the superior performance of *MobiAmbulance* over other representative comparison methods.

In the future, we plan to build models for predicting the appearance of patient pick-up requests by analyzing the human mobility under catastrophic situations. Future design and development activities will jointly consider the inferred flooding situation for scheduling of emergency vehicles, and emphasize a participatory design approach to incorporate users in the codesign process to make the *MobiAmbulance* more relevant and applicable to practice. Another scope for improvement is utilizing GPS data associated with emergency vehicle/ambulance travel during and following a disaster.

Acknowledgements

Thank Dr. Tho H. Nguyen at the University of Virginia, Dr. Mark W. Luckenbach and Dr. Lyle M. Varnell at the Virginia Institute of Marine Science in the College of William & Mary, and Dr. Xinyue Ye at the New Jersey Institute of Technology for their valuable discussion and advice. Thank X-Mode Social, Inc for providing the datasets. This research was supported in part by U.S. NSF grants NSF-1827674, CCF-1822965, OAC-1724845, ACI-1661378 and CNS-1733596, and Microsoft Research Faculty Fellowship 8300751.

References

[1] D. Berkoune, J. Renaud, M. Rekik, and A. Ruiz, "Transportation in disaster response operations," *Socio-Economic Planning Sciences*, vol. 46, no. 1, 2012.

[2] J. Holguín-Veras, M. Jaller, L. N. Van Wassenhove, N. Pérez, and T. Wachtendorf, "On the unique features of post-disaster humanitarian logistics," *Journal of Operations Management*, vol. 30, no. 7-8, 2012.

[3] Y. Yue, L. Marla, and R. Krishnan, "An efficient simulation-based approach to ambulance fleet allocation and dynamic redeployment," in *Proc. of AAAI*, 2012.

[4] V. Schmid and K. F. Doerner, "Ambulance location and relocation problems with time-dependent travel times," *European journal of operational research*, vol. 207, no. 3, 2010.

[5] S. Saisubramanian, P. Varakantham, and H. C. Lau, "Risk based optimization for improving emergency medical systems," in *Proc. of AAAI*, 2015.

[6] M. Gendreau, G. Laporte, and F. Semet, "A dynamic model and parallel tabu search heuristic for real-time ambulance relocation," *Parallel computing*, vol. 27, no. 12, 2001.

[7] D. Degel, L. Wiesche, S. Rachuba, and B. Werners, "Time-dependent ambulance allocation considering data-driven empirically required coverage," *Health care management science*, vol. 18, no. 4, 2015.

[8] P. L. Van Den Berg, J. T. Van Essen, and E. J. Harderwijk, "Comparison of static ambulance location models," in *Proc. of GOL*, 2016.

[9] L. V. Snyder and M. S. Daskin, "Reliability models for facility location: The expected failure cost case," *Transportation Science*, vol. 39, no. 3, 2005.

[10] M. S. Daskin, "A maximum expected covering location model: Formulation, properties and heuristic solution," *Transportation science*, vol. 17, no. 1, 1983.

[11] V. Schmid, "Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming," *European journal of operational research*, vol. 219, no. 3, 2012.

[12] M. S. Maxwell, M. Restrepo, S. G. Henderson, and H. Topaloglu, "Approximate dynamic programming for ambulance redeployment," *INFORMS Journal on Computing*, vol. 22, no. 2, 2010.

[13] C. Jagtenberg, S. Bhulai, and R. Van der Mei, "An efficient heuristic for real-time ambulance redeployment," *ORHC*, vol. 4, 2015.

[14] "X-mode social, inc," https://xmode.io/, accessed in April, 2019.

[15] "Openstreetmap," http://www.openstreetmap.org/, accessed in Jan, 2019.

[16] "Apache Hadoop," http://hadoop.apache.org/, accessed in July, 2018.

[17] "Apache spark," http://spark.apache.org/, accessed in July, 2018.

[18] Y. Zheng, "Trajectory data mining: an overview," *TIST*, 2015.

[19] D. Zhang, J. Huang, Y. Li, F. Zhang, C. Xu, and T. He, "Exploring human mobility with multi-source data at extremely large metropolitan scales," in *Proc. of MobiCom*, 2014.

[20] Y. Zhang and D. Brown, "Simulation optimization of police patrol districting plans using response surfaces," *Simulation*, vol. 90, no. 6, 2014.

[21] Y. Zhang and D. E. Brown, "Police patrol districting method and simulation evaluation using agent-based model & gis," *Security Informatics*, vol. 2, no. 1, 2013.

[22] K. Helsgaun, "An effective implementation of the lin–kernighan traveling salesman heuristic," *EJOR*, vol. 126, no. 1, 2000.

[23] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO - Simulation of Urban MObility," *IJASM*, vol. 5, no. 3&4, 2012.

[24] M. M. Morsy, J. L. Goodall, G. L. O'Neil, J. M. Sadler, D. Voce, G. Hassan, and C. Huxley, "A cloud-based flood warning system for forecasting impacts to transportation infrastructure systems," *Environmental modelling & software*, vol. 107, 2018.

[25] M. Hartswood, R. Procter, R. Slack, A. Voß, M. Büscher, M. Rouncefield, and P. Rouchy, "Co-realization: Toward a principled synthesis of ethnomethodology and participatory design," in *Resources, co-evolution and artifacts*.

[26] S. Khorshidi, M. Al Hasan, G. Mohler, and M. B. Short, "The role of graphlets in viral processes on networks," *Journal of Nonlinear Science*, 2018.

[27] S. Pohorec, M. Verlic, and M. Zorman, "Making applications more intelligent: A case study of optimizing route selection for emergency vehicles," in *Proc. of the 22nd IEEE International Symposium on Computer-Based Medical Systems*, 2009.