Meso- and Micro-scopic Routing of an Emergency Response Vehicle with Connected Vehicle Technologies

Gaby Joe Hannoun¹, Pamela Murray-Tuite², Kevin Heaslip³, and Thidapat Chantem⁴

¹Division of Engineering, New York University Abu Dhabi, UAE
²Civil Engineering Department, Clemson University, Clemson, SC 29634, USA
³Civil and Environmental Engineering Department, Virginia Tech, Blacksburg, VA 24061, USA
⁴Electrical and Computer Engineering Department, Virginia Tech, Arlington, VA 22203, USA

Abstract— Upon hearing the sirens and/or noticing the emergency lights, non-emergency response vehicles (non-ERVs) try to free space for the ERV. Yet, their combined actions do not always guarantee a safe and efficient passage. This paper introduces a semi-automated assistance system that facilitates the ERV movement in an urban transportation network by (1) determining the optimal ERV route from origin to destination, (2) identifying when and which non-ERVs to alert about the approaching ERV, (3) generating and sending, for each of these non-ERVs, a unique assistance message about how to appropriately react, and (4) disseminating intra-link maneuvering instructions to the ERV. The system benefits from real-time data from connected vehicles and relies on the execution of sequential integer linear programs (ILPs) as the ERV is moving towards its destination to limit the computation times as well as the impact on non-ERVs (non-ERVs not notified unnecessarily early). The proposed system accounts for links, intersections, and partial market penetration of connected vehicle technologies, making it applicable for short-term deployment. The system is evaluated using NetLogo, an agent-based modeling tool, in an urban transportation network with different combinations of congestion and market penetration levels. When compared to the case with no system (0% market penetration), results show ERV travel time reductions when the market penetration level is at a minimum of 40% (average of 9.09% travel time reduction at 100% market penetration) as well as a notable decrease in vehicular interactions (average of 35.46% and 81.38% reduction in ERV/non-ERV and non-ERV/non-ERV interactions, respectively, at 100% market penetration).

Keywords— Connected vehicles, Driver Assistance Systems, Emergency Services, Integer Linear Program, Intelligent Transportation Systems.

I. INTRODUCTION

Emergency response vehicles’ (ERVs) movements should be facilitated over their complete routes. Emergency preemption facilitates the ERV’s passage at intersections only and emergency warning systems such as lights and sirens are not fully effective [2]. The ERV movement’s support along transportation links received little research attention in the past. Hannoun et al. [1] introduced an integer linear program (ILP) that determines, using connected vehicle technologies, the ERV movement along a link segment while generating specific instruction messages to the group of non-ERVs expected to stop on this link segment. In [3], the system is first adjusted to accommodate partial market penetration levels (i.e., presence of unconnected downstream non-ERVs unable to share data or receive instructions) and, second, extended to allow the facilitation of ERV movements along larger transportation links using a sequential optimization approach. As the ERV is moving towards the destination and progressively receives maneuvering recommendations, new downstream vehicles have to be notified and provided with messages requesting them to stop at a position they can reach comfortably. The objective of this paper is to expand upon these prior works applicable to transportation links exclusively to achieve a complete network-wide system, hence facilitating the ERV intra-link movement between an origin/destination pair by instructing downstream traffic to move to specific locations prior to the ERV’s arrival. After identifying the optimal ERV route from origin to destination using the hyperstar routing algorithm [4], a criticality analysis module, introduced in this paper, determines when and which group of downstream non-ERVs to warn and provide with an instruction message. When a group of non-ERVs is identified, the ILP, introduced in [1], is executed with
adjustments to accommodate the presence of intersections. The system is evaluated on a sample transportation grid network under different initial conditions using the NetLogo agent-based modeling environment. Results in terms of ERV travel times and vehicle interactions are investigated and compared to a currently adopted practice for different market penetration levels.

The remainder of this paper is divided into eight sections. Section II includes a review of previous studies addressing emergency response vehicle operations. Next, the system’s general approach is presented (Section III) followed by a detailed description of each of its modules (Sections IV-VI). The experimental analysis plan and results are discussed in Section VII and Section VIII respectively. Finally, Section IX presents the key findings.

II. LITERATURE REVIEW

The majority of trauma fatalities occur before hospital care [5], making the efficiency of emergency response systems critical. The emerging capabilities of Intelligent Transportation Systems (ITS) technologies brought considerable benefits to urban transportation networks and to emergency response systems in particular. For example, automatic crash notification systems that mainly rely on the transmission of sensor data have resulted in reduced ERV response time and victim mortality [6-9].

Research has addressed different aspects of ERV movement. Signal pre-emption improves ERV movement through intersections [10-12] which are major conflict points where most of the ERV-related crashes occur [13]. Systems for dynamic dispatching and relocation of idle emergency vehicles have been developed [14-16] to improve the emergency response system’s coverage and response times. As for the routing of a single ERV from an origin to a destination, general (i.e., not ERV specific) dynamic shortest-path algorithms are usually employed [16].

Our system adopts Bell et al.’s time-dependent hyperstar routing algorithm [4] for a single ERV traveling through an urban transportation network. The selected algorithm [4] considers the dynamic aspect of urban transportation networks by assuming that link performance measures are not static. That is, the latter varies with time of day and are subject to uncertainty due to irregular incidents. The hyperstar routing algorithm generates, using a goal-directed search, a set of alternative routes, called a hyperpath, leaving the final route selection within the ERV driver’s discretion (further discussed in Section IV).

Within the overall ERV route, current techniques used by ERVs to facilitate their passage along streets are strobe lights and sirens [10]. These traditional warning systems do not fully eliminate confusion and risks of vehicle collisions [2]. Several researchers developed more advanced emergency warning systems [17-19] to provide surrounding non-ERVs with early warning messages and information to ensure that downstream traffic is well-aware of the incoming ERV with enough time and detail to react appropriately. However, the decision about how to react is left to the discretion of the drivers, which means that timely and proper actions are not guaranteed. Systems providing downstream traffic with early warning messages that also enclose advice about how to maneuver exist but are limited. Buchenscheit et al. [20] presented a conceptual prototype of an advanced emergency warning system that gives ERV route information and maneuver recommendations to downstream vehicles by relying on inter-vehic le communication. They evaluated the need for this additional safety measure using an expert survey which showed high acceptance. Inspired by Buchenscheit et al. [20], Weinert et al. [21] develop a system that creates a rescue lane for the ERV’s passage using V2V (vehicle-to-vehicle) communications. A static rule-based approach is adopted to generate the warning messages and determine whether the non-ERVs should change lanes. This approach assumes a predefined ERV intra-link path and that the non-ERVs that are travelling on the ERV’s lane are the only ones of interest (i.e., they need to move) [21]. However, a more dynamic intra-link ERV path that is generated based on the downstream traffic may lead to better travel time [1]. In other words, an ERV steering away from a platoon of vehicles may improve its travel time and lead to reduced non-ERV interactions than an ERV maintaining a pre-determined movement and expecting the downstream non-ERVs to cooperate and create a rescue lane. Potentially reducing vehicle interactions, Yoo et al.’s [22] system requests downstream non-ERVs to shift away from the lowest density lane to reserve it for ERV use. Yet, this system needs to be extended to account for desired lanes at upstream and downstream intersections, as the lowest density lane may not always be optimal and appropriate for ERV use (e.g., turning movements).

This paper extends a previously introduced system [1, 3] that relies on connected vehicle technologies to identify the optimal ERV intra-link movement using a mathematical program. This system [1, 3] does not assume a pre-defined ERV movement, but identifies it based on the downstream traffic by minimizing the ERV travel time and vehicular interactions. It also determines optimal maneuvering recommendations to the ERV as well as non-ERVs instead of following a set of static rules that may not be feasible (or safe). This system acts as a special type of lane change advisory [23-26] that grants priority to ERVs. The system
proposed in [3] adopts a sequential optimization approach that consists of executing mathematical programs consecutively as the ERV is traveling, hence enabling the optimization of the ERV movement along large transportation links while also considering the presence of unconnected vehicles. This paper significantly extends the system in [3] to take into consideration the desired lane occupancies prior to each signalized intersection depending on the movement to be executed at that intersection (based on the selected route to the destination) to ensure a safer trip. It also combines the sequential approach, previously introduced in [3], with other components such as a vehicle routing to present a complete application.

III. PROPOSED SYSTEM

The proposed centralized system collects and processes data from connected vehicles (ERV and non-ERVs) to find the ERV’s optimal intra-link route and progressively warn and disseminate instruction messages to connected non-ERVs while concurrently delineating and informing the ERV about the fastest and safest intra-link path to adopt. As the system aims to provide these different yet complementary functions, it is structured into three interrelated modules, as shown in Figure 1. First, the ERV route generation module determines the optimal route in the transportation network using the time-dependent hyperstar routing algorithm [4]. The major contributions of this paper are the following two modules which depend on the selected ERV route from origin to destination but not on the technique used to generate such a route. The second module is the Criticality Analysis Module (CAM) that screens the downstream non-ERVs as the ERV is travelling towards its destination to identify the vehicles that should be notified about the approaching ERV and informed about how to effectively react at the earliest time possible. It is only when a vehicle is identified by the CAM that the third module, called the Sequential Integer linear program Optimization Module (SIOM), is activated. This last module preprocesses the vehicle data collected from the connected non-ERVs identified by the CAM and executes an ILP, initially introduced in [1], and extended in this paper to account for desired lane occupancies at intersections, to ensure that the ERV can smoothly and safely follow the route selected in the first module. By maximizing the speed of the ERV and the free space around it, the mathematical program (1) optimizes the ERV’s intra-link movement along the downstream road segment where the critical non-ERVs can stop, (2) determines the best ERV maneuvering actions to take along that segment, and (3) identifies where each non-ERV should stop before the arrival of the ERV. The SIOM optimizes the intra-link movement of the ERV along short link segments on the selected ERV’s route sequentially as the ERV is travelling and as new vehicles come within range of the ERV’s movements. To ensure the feasibility of this system in real-time and to make it scalable to large networks and specifically long ERV trips, SIOM adopts this sequential approach as it directly controls the ILP’s computation times. One of the preprocessing steps in this module considers the output of previous ILP optimizations to ensure continuity of the ERV movement and consistency of the instruction messages being disseminated to the ERV and non-ERVs. The SIOM considers the presence of signalized intersections and infers based on the upcoming movement at the intersection, the desired lane to be occupied by the ERV on the upstream and downstream link of the intersection to ensure a smooth ERV passage. In addition, as the transition to a fully connected vehicle environment is expected to take decades [27], it is crucial to account for partial market penetration conditions in this proposed system. This third module applies an estimation technique, as a preprocessing step prior to solving the ILP, to predict the presence of unconnected non-ERVs between pairs of connected non-ERVs based on a distance criterion. These additional non-ERVs considered in the ILP virtually reserve extra spots for the actually present unconnected non-ERVs.

Data flows between the proposed centralized system, the ERV and non-ERVs are shown in Figure 1. The one-time data flow occurs once between entities. The periodic data flow occurs periodically while the “as-needed” data flow is a data exchange that occurs whenever new information or an update is available (i.e., when the ERV intra-link movements or the non-ERV instructions are generated). In this paper, it is assumed that all vehicle-to-infrastructure communications (i.e., ERV/non-ERV to proposed centralized system) are assured. Assessing the implications of imperfect communications is part of future work. Each of the modules is discussed thoroughly next.

IV. ERV ROUTE GENERATION MODULE

The expected arrival time’s reliability is a major concern for navigation system users. Generating a single path from an origin to a destination by excluding all unreliable links can result in a path that may still be unreliable possibly because it is not easy to drive (e.g., many turns). This is why it is important to consider the generation of a set of good alternative paths especially as congestion materializes with time and after a trip starts. The time-dependent hyperstar algorithm introduced by Bell, Trozzi [4] identifies a set of alternative optimal paths, called hyperpaths, from origin to destination by minimizing the expected arrival time to the
The hyperstar algorithm allows driver’s preferences, which is important for ERV drivers, who are familiar with the area they are serving, as they may select their preferred route based on their previous experience. Due to the large number of links in transportation road networks, efficient path finding algorithms are highly recommended and this algorithm incorporates a goal-directed search to accelerate the computation. The hyperstar algorithm is regarded as a hyperpath version of the time-dependent Astar algorithm, which is a speed up of Dijkstra’s classic shortest path algorithm. The time-dependent hyperstar algorithm uses two dynamic link performance measures: the undelayed travel time (varying with different times of day), and the maximum delay experienced due to unexpected events. This algorithm captures the dynamic aspect of road networks by assuming that link performance measures used for the path finding algorithm are not static; they vary with the time of the day and are subject to uncertainty due to irregular incidents, a key consideration for ERVs. Depending on the available data sources, the undelayed link travel time can be obtained from historic data, sensor data and/or data from connected vehicles. The maximum delay can be assigned to a link due to unexpected events such as vehicle crashes. This vehicle routing algorithm is considered a good fit to the proposed system. Yet, it can be replaced by any other routing algorithm.

V. CRITICALITY ANALYSIS MODULE (CAM)

As the ERV is moving, the CAM identifies the next set of downstream non-ERVs which should receive an instruction message. The CAM determines a dynamic detection range (discussed below) measured from the ERV’s current position. This range is defined to allow a detected vehicle enough time to react and execute the instruction, generated by SIOM, before the ERV’s arrival. The CAM identifies the non-ERVs whose real-time information should be sent to the SIOM after filtering the non-ERVs to avoid generating and sending unnecessary instruction messages to vehicles that are expected to diverge from the ERV’s route at intersections.
before the ERV’s arrival. The CAM is executed regularly as the ERV is moving towards its destination. The smaller the time step (i.e., CAM’s execution frequency) supported by the available wireless communication network, the smaller is the risk of not identifying a vehicle that needs to be notified on time.

A. CAM Methodology

CAM uses the following approach to identify the groups of non-ERVs below.

1) Non-ERVs of Interest: These are the connected non-ERVs that will potentially interact with the ERV if no instruction message is sent to them. They are downstream of the ERV and on its remaining route but have not yet received an instruction message from the system. Non-ERVs travelling on the opposite direction of the ERV route are not considered as contraflow operation is not in the scope of this paper. The CAM identifies the non-ERVs traveling along the links as well as the non-ERVs inside intersections heading to a link in the ERV’s route based on data about their compass direction. Based on data about emergency preemption detection ranges [28], the emergency preemption has a higher detection range than the system’s detection range which is discussed next. So, the CAM does not consider the vehicles that have not entered the ERV’s route yet as it assumes that emergency preemption at signalized intersections prohibits these non-ERVs from entering the intersection and hence joining the ERV’s route. New vehicles join the ERV route at intersections where preemption is not yet activated, but these non-ERVs are expected to be considered later, when the ERV gets closer to them.

2) Critical Non-ERVs: From the set of non-ERVs identified previously, the ones within a distance of ($\Delta d_{det}$ + $\varepsilon$) from the current ERV’s position are identified. The minimum detection range $\Delta d_{det}$ includes the critical non-ERVs (i.e., whose real-time information should be processed immediately so that the non-ERV receives the instruction and stops before the arrival of the ERV). A relatively short distance $\varepsilon$ is added to $\Delta d_{det}$ to include the non-ERVs that may become critical during the CAM execution’s time step. $\Delta d_{det}$ and $\varepsilon$ are computed using Equations (1) and (2) respectively that are developed based on time-space diagrams.

$$\Delta d_{det} = FS \times ((\sigma^e - \sigma^v) \times (tcc + t_r + \frac{\sigma^v}{g} + \frac{(\sigma^v)^2}{2g})$$

$$\varepsilon = \Delta t \times (\sigma^e - \sigma^v)$$

Where: ($\sigma^v$) is the average non-ERV speed (ft/sec); ($\sigma^e$) is the current ERV speed (ft/sec); ($t_r$) is the non-ERV’s reaction time (sec); ($tcc$) is the communication and computation time interval (sec), which is a function of the number of non-ERVs and the largest longitudinal distance between a pair non-ERVs; ($g$) is a comfortable non-ERV deceleration (ft/sec2); ($\Delta t$) is the time-step of the CAM; and ($FS$) is a factor of safety to account for additional time needed for non-ERVs to change lanes.

3) Additional Non-ERVs: The set of critical non-ERVs is expanded to add those that may become critical and within a distance of ($\Delta d_{det}$) during the time of communication and computation ($tcc$) of the next ILP optimization. These non-ERVs are located beyond the critical non-ERVs and within a distance equal to ($\Delta d_{det} + \varepsilon + \Delta d_{add1}$) downstream of the ERV, where $\Delta d_{add1}$ is determined using Equation (3). If more non-ERVs can be added without affecting the maximum allowable communication and computation time for the next ILP optimization, then the non-ERVs that are within a distance of ($\Delta d_{det} + \varepsilon + \Delta d_{add2}$) from the ERV’s current position are identified, where $\Delta d_{add2}$ is computed using Equation (4). These non-ERVs, if added to the group of non-ERVs for the next ILP optimization, will have an estimated waiting time at a final position that does not exceed a maximum tolerated duration ($\Delta t_{wait}$). The complete set of non-ERVs to undergo the next sequential ILP optimization and to receive instruction messages consists of the critical non-ERVs and additional non-ERVs groups.

$$\Delta d_{add1} = tcc \times (\sigma^e - \sigma^v)$$

$$\Delta d_{add2} = \sigma^e \times \Delta t_{wait}$$

B. Filtering of Vehicles

Due to the presence of intersections, some vehicles may diverge from the ERV’s route after the data collection. In addition, it is not recommended to ask a non-ERV to brake and reach a final position if its minimum stopping distance (i.e., minimum distance needed to reach a full stop) only allows it to stop beyond an intersection at which it was planning to exit the ERV’s route. This non-ERV is not required to stop or follow instructions generated by the system. To avoid unnecessary communication, filtering the non-ERVs determined in Section V.A. is required.

A parameter $\varphi_j$ is defined for each non-ERV $j$ and takes the value of 1 if $j$ is excluded from the next ILP optimization and 0 otherwise. If non-ERV $j$ is initially positioned on a link in the ERV’s route and upstream of an intersection and if its minimum final position ($MFP_j$) defined based on its minimum stopping distance ($MSD_j$) is beyond the intersection’s stop line and if it is expected to remain on the ERV’s route, then non-ERV $j$ should receive an instruction
the optimization of the emergency preemption of traffic signal control, it is assumed that emergency preemption at the next intersection is activated prior to the ERV departure’s time to allow the discharge of the stopped queue on the ERV approach so that the ERV does not approach non-ERVs that are still stopped at the intersection.

The proposed system sends instruction messages to downstream traffic. A vehicle that receives a message should follow the instruction and reach the assigned final position before the ERV’s arrival. This is why new vehicles are prohibited entry to the first link of the ERV’s route for a short time ($\Delta t_{ne}$) prior to the ERV’s desired departure time using traffic control upstream of the ERV’s origin. Consequently, at the desired time of ERV departure, the non-ERVs present on the first link are at the minimum detection distance of $(\Delta d_{det} + \delta)$ from the ERV’s initial position. $\Delta t_{ne}$ is computed using Equation 5:

$$\Delta t_{ne} = \frac{\Delta d_{det} + \delta}{\sigma}$$

**VI. SEQUENTIAL ILP OPTIMIZATION MODULE (SIOM)**

The SIOM is an extension to the approach presented in [1] that is applicable to a short transportation link without considering the presence of intersections at which non-ERVs can diverge from the ERV’s route. This modified approach relies on input data deduced from the ERV route generation module (discussed in Section IV) and CAM (discussed in Section V). In this section, the system is described with its aspects that remained unchanged and the new extensions that allows it to be applicable in an urban transportation network.

**A. System Description**

The ILP identifies the best ERV intra-link passage based on downstream traffic by maximizing the ERV speed and the free space adjacent to its movement. Based on the ILP’s output, the intra-link path and set of maneuvers for the ERV are determined and each downstream non-ERV is provided with a location along the link at which to stop. These outputs can be communicated to the appropriate vehicles through connected vehicle technologies. Note that the ERV speed integer variables are expressed in speed stages (see Appendix in [1] for more details) to avoid unrealistic acceleration and deceleration rates. As the ERV is travelling towards its destination, new non-ERVs become critical. The ILP should not be applied to optimize the positions of all non-ERVs that are present on the ERV’s remaining route at its time of departure and all at once because (1) this will engender high computation times making the system inefficient in real time and for large trips and (2) this will request non-ERVs to stop extremely early and unnecessarily
in case they were intending to diverge from the ERV’s route before its arrival. Subsequently, only the filtered set of critical and additional non-ERVs identified in CAM (discussed in Section V) are the ones considered in this module. Data about the position and speed of each of the non-ERVs is needed as input to the ILP to ensure that a non-ERV is assigned to a position it can safely reach. A preprocessing step estimates the presence of unconnected vehicles to accommodate partial market penetrations. Non-ERVs travel along an initial range (IR) at the time of data collection. The feasible stopping range (FSR) of each non-ERV is determined to identify the optimization range, called the assignment range (AR) along the link. Based on the output of this ILP, the ERV intra-link movement and non-ERV assignment positions along this AR are determined. Next, as the ERV is travelling, a new ILP execution is required with new downstream non-ERVs. The corresponding AR location along the link of this new ILP depends on the FSR of the new non-ERVs and may overlap or form a gap with the previous ARs, as shown in Figure 2. Therefore, the output of previous optimizations should be considered to ensure continuity of the ERV movement and to avoid any confusion related to assigning more than one non-ERV to the same location.

The ILP presented in [1] is applicable to a single transportation link segment. All detected non-ERVs are on the same transportation link and are expected to remain on it after data collection and until they complete their corresponding instruction. The optimization space (AR) does not include an intersection in which non-ERVs cannot stop and after which non-ERVs may diverge from the ERV’s route. Subsequently, the proposed system previously presented in [1] is adjusted for a transportation network with intersections. Now, non-ERVs may be on different transportation links of the ERV route at the time of detection or may be assigned to final locations on different links of the ERV route. Hence, non-ERVs may have different headings in the IR and/or in the AR, as shown in Figure 3. Heading is a number in degrees between 0 and 359.9, measured from the north in a clockwise direction (i.e., 0 is north, 90 is east). As presented in [1], the system discretizes a transportation link into identical cells. The x-axis represents forward (longitudinal) motion while the y-axis represents lateral motion. To ensure consistency of the system’s setup, the X and Y axes rotate with the different headings, as shown in Figure 4.

Non-ERVs cannot be assigned to positions inside the intersections. The ERV has a predefined movement inside the intersection that depends on the next link in its optimal route. Since the intra-link movement inside the intersection...

Figure 2 Initial ranges (IRs) and corresponding assignment ranges (ARs)

Figure 3 Different heading for different links upstream an intersection

Figure 4 X and Y axis for different headings
is fixed (i.e., does not need to be generated by the ILP) and since non-ERVs are not supposed to stop in the intersection area, there is no need to include the intersection space in the ARs (i.e., optimization spaces) and to unnecessarily increase the problem size and computation time. Figure 5-a through Figure 5-e describe how the ERV route in an sample transportation network (highlighted in Figure 5-a) is pictured by the SIOM: as a set of multiple links with the same or different headings and without intersections (Figure 5-e).

Since the sequential ILP optimization omits the intersection space, it virtually places the non-ERVs, that are in reality initially positioned inside the intersection at the time of detection, on the stop line prior to that intersection. Yet, the lateral position on the stop line would depend on the movement being performed inside the intersection. As a preprocessing step prior to the activation of the ILP, each non-ERV receives an ID that increases with the distance (along the X axis) to the ERV and with the lateral position (along the Y axis), regardless of whether it is positioned on a link or in the intersection. It is due to this ID that the ILP infers the true positions of the vehicles that may look overlapping. The ILP subsequently generates an output that ensures that there is no weaving among non-ERVs and that no more than one non-ERV is assigned to a cell along a link.

Depending on the optimal ERV route generated by the first module described in Section IV, the desired lateral positions upstream and downstream of each intersection in the ERV route are deduced based on the movement to be performed inside the intersection. For example, the ERV should occupy the leftmost lane before entering an intersection in which it will make a left movement. Table I and Table II present the notation of the variables and parameters used in the preprocessing steps and mathematical program described in Sections VI.B. and VI.C. respectively.

B. Preprocessing Steps

1) Unconnected Non-ERVs Presence: This system accommodates the presence of unconnected non-ERVs that are unable to share their real-time information and receive instruction messages when an ERV is approaching. Upon receiving the connected non-ERVs’ data, a preprocessing step estimates the presence of unconnected non-ERVs in the IR. The ILP is independent of the estimation technique. This technique, is inspired by Feng et al. [30] where unconnected non-ERVs are added between pairs of connected non-ERVs using a distance criterion based on the Wiedemann car-following model for the slow-down region. Constraints are added to the ILP to differentiate between a connected and an (estimated) unconnected non-ERV. While the former is expected to follow the received instruction message, the latter’s movement is estimated by assuming a “follow the leader” behavior using a constraint (Equation 49). Two binary parameters are needed for this constraint. The first parameter is $t_y$ which refers to the type of the non-ERV $j$ by taking the value of 1 for connected and 0 for unconnected. The second parameter is $l_e$ which is the ID of non-ERV $j$’s leader. For connected non-ERVs, $l_e$ can refer to the leader of non-ERV $j$ even though the constraint (Equation 49) will not be binding due to $t_y$ equal to 1. If the leader of non-ERV $j$ is not present in the same ILP optimization, $t_y$ is set to the ID of $j$.

2) Feasible Stopping Range (FSR) of each Non-ERVs: This step consists of determining, for each connected non-ERV $j$ (detected by the CAM) and unconnected non-ERV (added in the previous step), a FSR within which non-ERV $j$ can comfortably and safely stop [1.]. The $MSD_j$ of each non-ERV $j$ is determined using Equation 6 which includes (1) the distance travelled during the computation time ($t_{cc}$) and during the reaction time ($t_r$) and (2) the braking distance using a maximum comfortable deceleration of $(\delta_{max})$. The
(MSD) is computed in terms of cells and cannot be smaller than 1 cell. This is to make sure that a nearly stopped non-ERV does not receive an instruction to shift lanes without moving forward. After determining the (MSD) of each non-ERV j, the minimum final position (MFP) with respect to the start of the IR is identified using Equation 7.

\[
MSD_j = \max \left( \left( \frac{(tcc+\sigma_j)\sigma_j + s_j^2}{2s_j} \right) \right) \quad (6)
\]

\[
MFP_j = x_j^0 + MSD_j \quad (7)
\]

3) **AR and Minimum Final Index of each Non-ERV in the AR**: The ILP optimizes the ERV intra-link movement and non-ERVs final positions over a link segment that is called AR [1]. The AR is a range along the link segment that includes the FSR of all the non-ERVs in the current optimization. An AR is defined as to limit the problem size because there is no need to execute the optimization over a space which no non-ERV can utilize. The AR starts at least one increment prior the smallest MFP and ends at or after the largest (MFP + c) in a way to make the AR’s longitudinal size a multiple of the ERV’s longitudinal size +

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>Size of a cell, measured along the X axis (in ft)</td>
</tr>
<tr>
<td>W</td>
<td>Size of a cell, measured along the Y axis (in ft)</td>
</tr>
<tr>
<td>N</td>
<td>Size of an ERV, measured along the X axis (in cells)</td>
</tr>
<tr>
<td>LL</td>
<td>Size of the AR, measured along the X axis (in cells)</td>
</tr>
<tr>
<td>Y</td>
<td>Size of the AR (traversable shoulders included), measured along the Y axis (in cells)</td>
</tr>
<tr>
<td>j</td>
<td>Number of non-ERVs</td>
</tr>
<tr>
<td>ARstart</td>
<td>X index of the most upstream cell in the AR, with the start of the IR as origin</td>
</tr>
<tr>
<td>SL</td>
<td>X index of the cells forming the stop line of the next intersection, with the start of the AR as origin (in cells)</td>
</tr>
<tr>
<td>t (^r)</td>
<td>Reaction time of non-ERVs (in seconds)</td>
</tr>
<tr>
<td>(\sigma_j)</td>
<td>Initial speed of non-ERV j (in fps)</td>
</tr>
<tr>
<td>(\delta_j^{max})</td>
<td>Maximum comfortable deceleration rate of non-ERV j (in fps(^2))</td>
</tr>
<tr>
<td>tcc</td>
<td>Communication and ILP computation time (in seconds)</td>
</tr>
<tr>
<td>x(_j^0)</td>
<td>Initial index, along the X axis, of non-ERV j with the start of the IR as origin</td>
</tr>
<tr>
<td>y(_j^0)</td>
<td>Initial index, along the Y axis, of non-ERV j</td>
</tr>
<tr>
<td>MSD(_j)</td>
<td>Minimum stopping distance of non-ERV j, measured along the X axis (in cells)</td>
</tr>
<tr>
<td>MFP(_j)</td>
<td>Minimum final index along the X axis of non-ERV j with the start of the IR as origin</td>
</tr>
<tr>
<td>x(_j^f)</td>
<td>Minimum final index along the X axis of non-ERV j with the start of the AR as origin</td>
</tr>
<tr>
<td>c</td>
<td>Size of the FSR measured along the X axis, also called FSR cutoff value (in cells)</td>
</tr>
<tr>
<td>(S_{min})</td>
<td>ERV minimum speed (in speed stage)</td>
</tr>
<tr>
<td>(S/ree)</td>
<td>ERV maximum speed (in speed stage)</td>
</tr>
<tr>
<td>(np_{x,y})</td>
<td>Binary parameter that takes the value of 1 if cell (x,y) is already assigned to a non-ERV from a previous ILP optimization and 0 otherwise</td>
</tr>
<tr>
<td>(\partial_{k}^{x,y})</td>
<td>Binary parameter that takes the value of 1 if instruction k was previously sent to the ERV at increment i and lateral index y and 0 otherwise</td>
</tr>
<tr>
<td>(su_{i}^{l})</td>
<td>Integer parameter that takes the value of (s_{c,mp}^{l}) if this variable was previously generated at increment i due to overlap with previous ARs and ((S/ree + 1)) otherwise (see Table II for (s_{c,mp}^{l}) description)</td>
</tr>
<tr>
<td>(\gamma_i)</td>
<td>Desired ERV lateral position</td>
</tr>
<tr>
<td>(\xi_i^{l})</td>
<td>Binary parameter that takes the value of 0 if (SL \geq (N + 1)i) and (SL &lt; (N + 1)(i + 1)) and 1 otherwise, for (i = 1, \ldots, LL/(N + 1) - 1)</td>
</tr>
<tr>
<td>(\xi_i^{f})</td>
<td>Binary parameter that takes the value of 0 if an ERV lateral position is desired at the end of the AR due to an upcoming intersection (i.e., if (LL \leq SL \leq LL + (Y - 2)(N + 1)i) and 1 otherwise</td>
</tr>
<tr>
<td>ty(_j)</td>
<td>Binary parameter equal to 1 if non-ERV j is a connected vehicle and 0 otherwise</td>
</tr>
<tr>
<td>(le_j)</td>
<td>Integer parameter equal to the label of the leader of each non-ERV j in the IR and equal to j if the leader of non-ERV j is not included in the IR</td>
</tr>
</tbody>
</table>
buffer. After defining the start of the AR (\(AR_{\text{start}}\)) with respect to the start of the IR, the minimum final index (\(x^f\)) of each non-ERV in the AR is determined, using Equation 8.

\[ x^f = \text{MFP}_j - AR_{\text{start}} + 1 \]  

(8)

4) **Dynamic FSR Cut-off Input Value \(c\):** The FSR of each non-ERV starts at its \(x^f\) and extends, along the X axis, a number of \(c\) cells beyond it. This is to limit the problem size and to make the computation time smaller by reducing the space within which each non-ERV can stop. The \(c\) parameter is an input to the ILP and is the same for all non-ERVs (this can be relaxed in the future). When no feasible solution is found, re-executing the problem with a higher \(c\) may lead to feasibility. In other words, as \(c\) increases, new combinations of non-ERVs’ final positions are added, potentially leading to a combination satisfying all constraints, yet resulting in higher computation times. The ILP computation time includes the duration spent while increasing \(c\) until a feasible solution is found. Subsequently, two rules of thumb are used to determine the minimum \(c\) value that should be used initially, to avoid iterations. First, the initial \(c\) value is increased until no FSR ends at or before the most downstream non-ERV reaches the desired position. Passing among non-ERVs in different ILP optimizations is prohibited. So, the non-ERVs in the current optimization should stop after the most downstream non-ERV in previous ARs. Second, the minimum \(c\) value should ensure that the required number of cells for ERV movement is provided. The ERV needs a minimum number of free cells to move along a link segment. If lane changes are performed, additional free cells are required to be able to make the maneuvers. The available number of cells is the difference between the total number of cells and the number of cells utilized by non-ERVs.

5) **ERV Speeds and Desired Lateral Position at Intersections:** As discussed in Section VI.A., the intersection space is not considered in the AR. If the AR encloses an intersection (i.e., starts at or before an intersection and ends beyond it), a given cell in the AR will either be part of the link upstream or part of the one downstream of the intersection. The most downstream cells of the link upstream of the intersection are the ones forming the stop line of the intersection. The ILP ensures that the ERV enters and exits the intersection from the appropriate lateral position depending on the intended movement inside the intersection. In addition, it is assumed that the ERV speed should decrease prior to an intersection and cannot increase inside the intersection. If the AR ends within \((Y - 2)\) increments prior to the stop line of an intersection, the ILP makes sure that the ERV leaves the AR at the desired lateral position that depends on its upcoming movement. Otherwise, the ERV will not have enough distance to make the lane changes required to enter the intersection from the appropriate lane. To account for the ERV speed implications and desired ERV lateral position due to intersections, new constraints

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w_{x,y}^i)</td>
<td>ERV assignment binary variable taking the value of 1 if cell ((x,y)) is part of the ERV path and 0 otherwise (where (x = 1, \ldots, LL) and (y = 1, \ldots, Y))</td>
</tr>
<tr>
<td>(s^i)</td>
<td>ERV speed integer variable equal to the ERV speed generated at increment (i) (where (i = 1, \ldots, LL/(N + 1)))</td>
</tr>
<tr>
<td>(d_k^{iy}^j)</td>
<td>ERV instruction binary variable taking the value of 1 if instruction (k) is sent to the ERV at increment (i) and lateral position (y) (where (i = 1, \ldots, LL/(N + 1) - 1; y = 1, \ldots, Y) and (k = 1) refers to move right, (k = 2) to straight and (k = 3) to move left)</td>
</tr>
<tr>
<td>(v_{x,y}^{j,y})</td>
<td>Non-ERV assignment binary variable taking the value of 1 if non-ERV (j) is assigned to cell ((x, y)) and 0 otherwise (where (x = 1, \ldots, LL; y = 1, \ldots, Y) and (j = 1, \ldots, J))</td>
</tr>
<tr>
<td>(s_{env}^i)</td>
<td>ERV speed environment integer variable equal to the ERV speed generated based on the ERV’s surroundings at every increment (i) (where (i = 2, \ldots, LL/(N + 1)))</td>
</tr>
<tr>
<td>(s_{\text{temp}}^i)</td>
<td>ERV temporary speed integer variable equal to the ERV speed generated at increment (i) based on the previous ERV instruction and ERV’s surroundings without accounting for the minimum and maximum allowable ERV speeds (where (i = 2, \ldots, LL/(N + 1)))</td>
</tr>
<tr>
<td>(\nu^i)</td>
<td>Binary variable taking the value of 1 at increment (i) when (s_{\text{temp}}^i \geq S_{\text{min}}) and 0 otherwise (where (i = 2, \ldots, LL/(N + 1)))</td>
</tr>
<tr>
<td>(p_{ij}^f)</td>
<td>Binary variable taking the value of 1 if no passing is occurring between non-ERVs (j) and (j') and 0 otherwise.</td>
</tr>
</tbody>
</table>
(Equations 50, 51 and 52) are added to the ILP with new parameters to be defined in this preprocessing step. The $S_L$, $\xi_1$ and $\xi_2$ are identified as defined in Table I. $\hat{y}$ takes the value of the desired ERV lateral position that depends on the movement to be performed inside the intersection and on the link composition.

6) Input from Output of Previous Optimizations: As a sequential approach is adopted, the output of the previous optimization sets the initial conditions of the next ILP which are (1) ERV initial speed and lateral position, (2) previously disseminated ERV instructions, (3) previously generated ERV speed variables, and (4) the most downstream non-ERV final position in previous ARs. The AR of the next ILP can overlap with a previous AR or start after the last AR forming a gap. This step should be executed after the identification of the next AR’s position.

First, as the ERV is travelling toward the destination, it may end up travelling at speeds that are different from the ones generated by the system to optimize the intra-link path. The ERV is equipped with connected vehicle technologies so the proposed system can track the ERV’s actual speeds. When a difference between the two speeds exists, the ERV speed stages at the increments that the ERV did not reach yet are adjusted to better estimate the ERV initial speed in the next ILP. In the case of AR overlap, the ERV initial lateral position and speed are deduced by retrieving those that were generated from previous optimizations at the increment that coincides with the first increment in the next AR. In case of a gap between ARs, the ERV is assumed to maintain a straight movement (same lateral position) and to increase its speed stage linearly up to $S_{free}$. In case an intersection exists over that gap and/or in the first ($Y-1$) increments of the next AR, a lane change is expected to be performed inside the gap and depends on the movement inside the intersection. In addition, a speed reduction prior to the intersection is expected because it is assumed that the ERV decreases its speed before entering the intersections located in this gap. Based on these assumptions, the ERV’s initial speed and initial lateral position of the ERV are deduced for the next ILP.

Second, in case of overlap, some increments may belong to more than one AR so it important to keep the instructions consistent. The instructions that were previously disseminated to the ERV should be left unchanged in future optimizations to avoid confusion. The most downstream ERV instruction that is disseminated to the ERV is the one delineating the ERV path up to the most downstream non-ERV final position in previous ARs. A binary parameter $(\delta^k_{\xi})$ takes the value of 1 if instruction $k$ was previously disseminated to the ERV at increment $i$ and lateral position $y$ and 0 otherwise or when no overlap exists.

Third, in case of overlap, the ERV speed at each increment in the overlap in the next ILP should be bounded by an upper limit (su1) equal to the ERV speed environment ($S_{env}^i$) that was previously generated at that increment, if any. The ERV speed environment is the ERV speed that takes only into consideration the presence of non-ERVs surrounding the ERV’s next movement. An upper limit (su1) is needed to account for the presence of non-ERVs from previous ARs in the next ILP, in the case of overlap. If an increment $i$ in the AR of the current optimization does not overlap with any increment of ARs of previous optimizations, then su1 takes the value of ($S_{free}^{i+1}$) because no ERV speed environment ($S_{env}^i$) was previously generated at that increment $i$.

Fourth, since no passing is allowed between non-ERVs considered in different ILP optimizations, the non-ERV with the most downstream final position and which has not yet diverged from the ERV route should be identified. This is to make sure that the non-ERVs in the next ILP are only instructed to stop at positions beyond it. The binary parameter $np^{x,y}$ takes the value of 1 at and before the most downstream non-ERV final position in previous ARs and 0 otherwise or when no overlap exists.

C. ILP Formulation

The objective function of this ILP maximizes the ERV’s speed at each increment (first component multiplied by $\alpha_1$) and the free space surrounding its intra-link movement (second component multiplied by $\alpha_2$). The ($S_{env}^i$) is the speed environment at increment $i$ that is only constrained by the number of non-ERVs on cells adjacent to the ERV’s movement prior to that increment (see Equations 32 through 37). Maximizing ($S_{env}^i$) ensures that the non-ERVs are pushed away from the ERV movement as much as possible even when the ERV is performing a right or left lane change prior to increment $i$ (i.e., when ERV speed ($s^i$) is lower than ($s^{i-1}$) due to the lane change) (see Equations 40, 45 and 46). The weight factors ($\alpha_1$) and ($\alpha_2$) are equal to 1 since this combination is the most unbiased [1]. Adopting an objective function that is composed of the first two components only may lead to several alternative solutions. The third component (multiplied by $\alpha_3$ that is a very small factor relative to $\alpha_1$ and $\alpha_2$) is added to favor, out of these alternative solutions, the ones with the most upstream non-ERVs final positions. The reason behind this preference is to better utilize the downstream space.

Maximize 
\[ z = \alpha_1 \sum_{i=2}^{N+1} s^i \ + \alpha_2 \sum_{i=2}^{N+1} s_{env}^i - \alpha_3 \sum_{x=1}^{X} \sum_{y=1}^{Y} v^{x,y,i}_f \]  
(9)
The constraints are as follows.

A cell can be occupied by a single vehicle (ERV or non-ERV),

\[ w_{x,y} + \sum_{j=1}^{s} v_{x,y} = 1 \quad \forall x; \forall y \tag{10} \]

A non-ERV should be assigned to only one cell in its FSR

\[ \sum_{x=x_{j}}^{x_{j}+c} \sum_{y=1}^{v_{x,y}} = 1; \forall j \tag{11} \]

\[ \sum_{x=x_{j}}^{x_{j}+c} \sum_{y=1}^{v_{x,y}} = 1; \forall j \tag{12} \]

Passing among all non-ERVs was prohibited in [1]. In this paper, this constraint is relaxed since prohibiting passing between all non-ERVs may be too conservative in some cases. For instance, no interaction is expected when two vehicles with different lateral positions are passing each other and maintaining their respective lanes. When dealing with varying speeds, a vehicle \(j\) initially positioned upstream of \(j'\) may have a downstream FSR than the one of \(j'\). If \(j\) is allowed to pass \(j'\), no interaction is possible as long as weaving between \(j\) and \(j'\) does not occur and \(j\) and \(j'\) are not assigned to the same lane. Equations 13 and 14 detect when two non-ERVs will pass each other to reach their generated final positions. They set the value of a binary variable \(p_{ij}'\) it is equal to 1 if no passing occurs between \(j\) and \(j'\) and equal to 0 otherwise. This parameter is then used in the weaving constraints (discussed next in Equations 15 and 16).

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} \leq \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} + LL(1 - p_{ij}') \quad \forall j' > j \tag{13} \]

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} + LLp_{ij}' \geq \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} \quad \forall j' > j \tag{14} \]

Weaving among non-ERVs is prohibited. In addition, when passing between \(j\) and \(j'\) will occur (i.e., \(p_{ij}' = 0\)), then \(j\) and \(j'\) are not allowed to stop on the same lateral position. For example, if passing will occur between \(j\) and \(j'\), then \(p_{ij}' = 0\). If \(j'\) is initially to the right of \(j\) (i.e., \(y_{j} \geq y_{j}'\)), then Equation 15 applies and \(j\) cannot be assigned to the same lateral position as the one of \(j'\). It can only be assigned to a lateral position strictly greater than the one of \(j'\). If passing will not occur between \(j\) and \(j'\), then \(p_{ij}' = 1\). If \(j'\) is initially to the left of \(j\) (i.e., \(y_{j} \leq y_{j}'\)), then Equation 15 applies and \(j\) can be assigned to the same lateral position as the one of \(j'\).

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} \leq \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} - 1 + p_{ij} \quad \forall j' > j \leq y_{j} \tag{15} \]

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} \geq \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} x \times v_{x,y} + 1 - p_{ij} \quad \forall j' > j \geq y_{j} \tag{16} \]

A minimum of one cell should be provided for the ERV at each longitudinal position.

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} v_{x,y} \leq Y - 1; \forall x \tag{17} \]

The ERV can receive only one instruction at each increment.

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} v_{x,y} \leq 1; \forall j \tag{18} \]

Go right and go left instructions cannot be generated at the rightmost lane and leftmost lanes, respectively.

\[ d_{1}^{j} = 0; \forall i = 1, ..., LL/(N + 1) - 1 \tag{19} \]

\[ d_{2}^{j} = 0; \forall i = 1, ..., LL/(N + 1) - 1 \tag{20} \]

If the current AR overlaps with previous ARs, the previously disseminated ERV instruction should be left unchanged, as discussed in Section VI.B.

\[ d_{3}^{j} \geq d_{k}^{j}; \forall k; \forall i = 1, ..., LL/(N + 1) - 1; \forall y \tag{21} \]

The ERV instruction and ERV assignment variables should be interconnected to ensure the continuity of the longitudinal and lateral ERV movement. When a given instruction is generated at an increment, the corresponding cells should be added to the ERV intra-link path.

\[ \sum_{y=1}^{v_{x,y}} w_{x,y} = 1; \forall x \tag{22} \]

\[ d_{1} \leq \frac{w_{i/(N+1)} + w_{i/(N+1)+1}}{2}; \forall i = 1, ..., LL/(N + 1) - 1; \forall t = 1, 2, ..., N + 1; \forall y > 1 \tag{23} \]

\[ d_{1} \leq \frac{w_{i/(N+1)} + w_{i/(N+1)+1}}{2}; \forall i = 1, ..., LL/(N + 1) - 1; \forall t = 1, 2, ..., N + 1; \forall y \geq 1 \tag{24} \]

\[ d_{1} \leq \frac{w_{i/(N+1)} + w_{i/(N+1)+1}}{2}; \forall i = 1, ..., LL/(N + 1) - 1; \forall t = 1, 2, ..., N + 1; \forall y \leq Y \tag{27} \]

The ERV instruction and non-ERV assignment variables should also be interconnected to ensure that the non-ERVs are not positioned on cells that are part of the ERV path or on the additional cells needed to perform right or left instructions.

\[ \sum_{y=1}^{v_{x,y}} \sum_{x=x_{j}}^{x_{j}+c} v_{x,y} \leq (N + 1)(1 - d_{2}^{j}); \forall i = 1, ..., LL/(N + 1) - 1; \forall y \geq 1 \tag{29} \]
The ERV speed environment is constrained by the presence of non-ERVs around the ERV’s next movement only. If the next movement does not have any non-ERVs stopped on its adjacent cells, then based on the surroundings, the ERV’s speed stage can increase by 1. In case of overlap, non-ERVs can only be assigned beyond the most downstream non-ERV’s final position in previous ARs. Thus, the speed environment variables along the overlap in the current AR consider the presence of non-ERVs in previous ARs. This is performed by limiting the ERV speed environment in the current ILP to speed upper limits determined in Section VI.B. 

\[
\sum_{i=1}^{\bar{t}} \sum_{j=1}^{I} v_{ij}^{xy} + \sum_{i=1}^{\bar{t}} \sum_{j=1}^{I} v_{ij}^{xyj-1} \leq (2N + 1)(1 - d_j^{xy});
\]

\[
\forall y > 1; \forall i = 1, ..., \frac{N}{N+1} - 1; t = (N + 1)i + 1;
\]

\[
\bar{t} = (N + 1)i + (N + 1); \bar{I} = (N + 1)i + N (30)
\]

\[
\sum_{i=1}^{\bar{t}} \sum_{j=1}^{I} v_{ij}^{xy} + \sum_{i=1}^{\bar{t}} \sum_{j=1}^{I} v_{ij}^{xyj+1} \leq (2N + 1)(1 - d_j^{xy});
\]

\[
\forall y < Y; \forall i = 1, ..., \frac{N}{N+1} - 1; t = (N + 1)i + 1;
\]

\[
\bar{t} = (N + 1)i + (N + 1); \bar{I} = (N + 1)i + N (31)
\]

The ERV speed at increment \(i\) should be equal to the maximum of \(S_{\text{env}}^{\text{min}}\) and \(s_{\text{temp}}^{\text{max}}\) and cannot increase beyond the maximum allowable speed \(S_{\text{free}}^{\text{max}}\).

\[
s^{i+1} \geq s_{\text{env}}^{\text{max}}, \forall i = 1, ..., LL/(N + 1) - 1
\]

\[
s^{i+1} \geq s_{\text{temp}}^{\text{max}}, \forall i = 1, ..., LL/(N + 1) - 1
\]

\[
s^{i+1} \leq S_{\text{free}}^{\text{max}}(1 - v^{i+1}) + s_{\text{temp}}^{i+1}, \forall i = 1, ..., LL/(N + 1) - 1
\]

\[
s^{i+1} \leq (S_{\text{free}}^{\text{max}} - S_{\text{env}}^{\text{max}}) v^{i+1} + S_{\text{env}}^{\text{max}}, \forall i = 1, ..., LL/(N + 1) - 1
\]

\[
\sum_{x=1}^{N} v_{xj}^{xy} + v_{ij}^{xy} \geq v_{ij}^{xyj}; \forall j; \forall x = 2, ..., LL; \forall y
\]

A final ERV lateral position should be imposed if an intersection exists at the end of the AR or within \((Y - 2)\) increments of its end.

\[
w^{(LL,y)} + \xi^{L} \geq 1
\]

If the stop line of an intersection is inside the AR, then, a lateral position is to be imposed before and after the intersection depending on the future movement of the ERV inside the intersection. \(\bar{Y}\) is the desired lateral position that depends on the ERV’s movement inside the intersection. Equation 51 sets the desired lateral position before entering the intersection and Equation 52 sets the desired lateral position when exiting the intersection (i.e., entering the downstream link). It is assumed that the same lateral position is maintained when exiting the intersection. This is why a “go straight” instruction will not be disseminated to the ERV to avoid confusion.

\[
w^{(N+1)(y)} + \xi^{L} \geq 1; \forall i = 1, ..., LL/(N + 1) - 1
\]

\[
w^{(N+1)(y)} + \xi^{L} \geq 1; \forall i = 1, ..., LL/(N + 1) - 1
\]
VII. Experimental Analysis

The proposed system is modeled in a simulation environment to assess its impact on traffic operations in an urban transportation network with signalized intersections. A 4 by 4 grid transportation network is modeled in NetLogo, an agent-based modeling language and environment. NetLogo is a widely used and open source tool that has already proved to be able to support microscopic traffic simulations [31]. The ability to implement the proposed complex system while visualizing the intra-link behavior of each vehicle motivated us to use this tool for experimental analysis [31]. Agent-based models are practical when simulating simultaneous operations and interactions of multiple agents. In our case, each vehicle (ERV or non-ERV) is an agent (called a turtle) that moves around in the NetLogo world. It reacts to its surroundings while aiming to follow any real-time message it may receive if equipped with connected vehicle technologies. Patches are the agents in NetLogo that form the ground over which the vehicles move [32]. This discrete aspect of the NetLogo world (space) fits the proposed system which also discretizes transportation links into identical cells with a size equal to that of a vehicle plus buffer. Using APIs, the model in NetLogo is invoked and controlled by a program running on a Java virtual machine. This java program activates the hyperstar routing algorithm when the ERV route generation module is triggered. The CAM, coded in NetLogo, periodically scans and filters the non-ERVs of interest. When a critical group of non-ERVs is detected, the sequential ILP optimization module, coded in a separate java program, is executed. The latter preprocesses the connected vehicles data from NetLogo and initiates the ILP that is coded in AMPL and solved using CPLEX. All tests are run on a computer with 3.1 GHz Intel Core i5 and 8 GB 2133 MHz LPDDR3 memory.

Due to the presence of signalized intersections in the system, emergency preemption should be considered. In real-life implementation of this system, any traditional or advanced emergency preemption technique can be used at signalized intersections. In this experimental analysis, a dynamic ERV detection distance is adopted for emergency preemption [11, 33]. The ERV detection distance depends on the time needed to switch the green indication to the ERV’s approach, the average queue length on the approach, and the ERV operating speed. Sensing technologies such as pavement inductive loops or microwave radars [11, 33] are assumed present for the estimation of the queue length. So, the dynamic detection distance is assumed to be insensitive to the market penetration level. This will allow a better assessment of the benefits of the proposed intra-link emergency assistance since these benefits are isolated from those due to adequate preemption activation, when compared to situations with zero market penetrations. Signals timing plans can be fixed (pre-timed), actuated, or adaptive [11]. Since emergency preemption is implemented at signalized intersections, it is expected to interrupt the cycle when activated, so the differences between the various types of signal timing plans can be considered minimal. In addition, since the focus of this system is the intra-link traffic operations and for simplicity, a pre-timed plan is adopted at intersections and maintained for all tests to ensure consistent comparison of the results. In addition, it is assumed that the connected vehicle needs a 2.5 seconds reaction time [34] before starting to brake and to move towards its assigned position. While an unconnected vehicle which cannot be identified by the CAM nor provided with a warning/instruction message, only detects the approaching ERV when located within the latter’s siren range, which is set to 555 ft [35].

The proposed system is tested under three v/c ratios of 0.75, 1 and 1.25, and different levels of market penetrations ranging between 0 and 100% at 20% increments. For the experiments conducted here, each link is composed of 3 lanes and 1 traversable shoulder. For each combination of market penetration level and congestion level, the results are compared to the case with no system (i.e., at 0% market penetration level). The performance measures used to assess the benefits of the proposed system are the ERV travel time and vehicular interactions (ERV/non-ERV and non-ERV/non-ERV). To quantify the impact of the proposed system, the average time during which non-ERVs were affected by the approaching ERV is recorded (discussed further in Section VIII.C.).

VIII. Discussion of Results

As discussed in Section VII, the system is executed on an urban transportation network with signalized intersections. A Poisson distribution is used to model traffic on each link in the network with average flow rates for each v/c ratio based on [36]. After the seed period, traffic generation will stop as vehicles leaving the network (i.e., NetLogo world) will reenter it from the opposite edge because the NetLogo world topology wraps vertically and horizontally (i.e., opposite sides are actually connected). This will emulate the presence of other intersections outside the network we modeled. For each combination of v/c ratio and market penetration level, 5 runs were initially conducted. Then, based on the average ERV travel time and standard deviation, the minimum sample size for each combination is identified using a 90% confidence interval and a 5-second
desired margin of error. Accordingly, more runs are executed until the minimum sample size is satisfied.

A. ERV Travel Time Benefits

The reduction in ERV travel time is the main indication of the effectiveness of an emergency assistance system. Figure 6a through Figure 6c show the variation in percent reduction in ERV travel time with different market penetration levels, when compared to the no-system scenario with 0% market penetration, for v/c ratios of 0.75, 1 and 1.25 respectively. Different v/c ratios just reflect different flow rates used to populate the network during the seed period. We acknowledge that a single v/c ratio cannot represent the state of a complete network. The proposed system succeeds in improving ERV travel time at all v/c ratios when the market penetration level is equal to or higher than 40%. As shown in Figure 6-b for a v/c ratio equal to 1, an increase in ERV travel time is noted, which makes the proposed system not recommended at 20% market penetration although significant reductions in vehicular interactions potentially result (discussed in Section VIII.B.). For all v/c ratios, the observed percent reduction in ERV travel time increases as the market penetration level increases. This increasing trend is expected. With a higher proportion of connected non-ERVs on the link, a faster passage for the ERV is more achievable. With partial market penetration, higher risks of interactions between the ERV and non-ERVs exist (discussed in Section VIII.B.), hence negatively affecting the ERV’s movement. Unconnected non-ERVs stopped on the nearest edge may force an ERV that was travelling on the edge to deviate to another free lane, resulting in a speed decrease. In addition, unconnected non-ERVs may occupy a position at which a connected non-ERV is instructed to stop. Hence, this connected non-ERV may end up impacting the ERV movement by stopping in a cell adjacent to the ERV’s path. At higher market penetration, the impact of unconnected vehicles on the ERV travel is minimized. At 100% market penetration, the highest percent reductions in ERV travel time are observed since all non-ERVs are moving to their assigned positions that are not at risk of

Figure 6 Percent reduction in ERV travel time for each market penetration level at a v/c ratio equal to 0.75 in (a), 1 in (b) and 1.25 in (c)
being occupied by other unconnected non-ERVs. It is important to note that this system may not offer tremendous benefits in terms of ERV travel time (range between 1% and 11%) as the ILP is generating the ERV passage while setting very restrictive rules about the weaving interactions between non-ERVs. The system may instruct non-ERVs to stop at locations adjacent to the ERV lane if this is the only position that would satisfy the no-weaving rule. If this is too conservative, the corresponding constraints in the ILP can be relaxed. In addition, the ERV optimal movement may include a right/ left lane change, that results in a speed decrease, just to avoid having non-ERVs stopped adjacent to its movement. If such a lane change is perceived as unnecessary (more weight can be attributed to the first component in the objective function compared to the second one, i.e., $\alpha_1 > \alpha_2$ can be used). The next subsection investigates the benefits of this system in terms of vehicular interactions.

**B. Vehicular Interactions**

Limiting vehicular interactions among non-ERVs and between the ERV and non-ERVs is a major benefit offered by this system. Making sure that the ERV can move forward safely is paramount but ensuring that non-ERVs are also safe is what makes this system valuable. Each non-ERV is assigned to a position that is at or beyond its minimum stopping distance, while limiting weaving and passing with others. Interactions among non-ERVs is assessed by counting risky maneuvers between non-ERVs. A risky interaction between two non-ERVs is recorded when both are aiming towards the same lane and one non-ERV had to pass the other. On the other hand, the system generates recommendation messages that try to direct the non-ERVs away from the ERV’s movement (by maximizing the second component in the objective function). To assess the amount of ERV/non-ERV interactions, the number of non-ERVs adjacent to the ERV’s movement are counted for each scenario. Figure 7 shows the average percent reduction for each of the non-ERV/non-ERV and ERV/non-ERV interactions for different market penetration levels. A reduction is observed at all market penetration levels for both types of vehicular interactions along with an increasing trend as the market penetration level is higher, hence confirming that this system enhances the safety of vehicles. Higher reductions in non-ERV/non-ERV interactions are observed when compared to the one between an ERV and non-ERVs. This is expected because as mentioned before, a feasible solution generated by the ILP may assign the non-ERVs to positions that are adjacent to the ERV path only to satisfy the no-weaving constraint.

**C. Impact of the Proposed System and Limitations**

From the non-ERVs’ perspective, an ideal emergency assistance system notifies them only when it is needed, in a way to grant them enough time to move and create the optimal passage for the ERV while most importantly minimizing the time they have to wait for the ERV at their final position. A system sending unnecessary recommendation messages to non-ERVs and/or requesting them to wait for a very long duration before the ERV’s passage is not effective and will not be embraced. By considering the non-ERVs within a detection range in the CAM (Section V.A) and then filtering out the non-ERVs that are diverging from the ERV’s route before its arrival (Section V.B.), the system identified the non-ERVs of interest which should be provided with an instruction

![Figure 7 Percent reduction in vehicular interactions for different market penetration levels](image-url)
message. To assess the impact of the proposed system on non-ERVs, the average time during which non-ERVs were affected by the ERV’s presence is recorded at each market penetration level. This duration starts upon receiving the instruction message and ends when the ERV becomes downstream. Note that non-ERVs’ recovery after the ERV’s passage is not investigated in this paper and will be included in future works. In Table III, the average percent increase in impact on non-ERVs (in terms of duration affected by the ERV’s passage) when compared to the case at 0% market penetration, is shown for each market penetration level. The system did not offer a reduction in ERV travel time at a 20% market penetration which means that the system was ineffective, hence the lowest impact on the non-ERVs. At market penetrations higher than 20%, the average percent increase in impact on non-ERVs decreases as the market penetration become higher.

The detection distance ($\Delta d_{det}$) includes a factor of safety ($FS$) primarily to account for the additional time required by non-ERVs to change lanes. In all the tests executed above, $FS$ is set to 1.25. An increase in $FS$ will directly result in higher detection distances, which should potentially lead to better ERV travel times because non-ERVs have more time to react and reach their assigned positions. In other words, due to possible inaccuracies of the heuristic approach adopted in the CAM (such as the use of an expected ILP computation time), the estimated duration that a non-ERV needs to reach its final position may be underestimated. To confirm this trade-off, a set of additional tests are executed with a factor of safety equal to 1.5 at 100% market penetration. The average reduction in ERV travel time increased from 9.09% (with $FS=1.25$) to 15.40% (with $FS=1.5$) while the average increase in non-ERV impact increased from 7.30% to 11.60%. These results are expected; increasing the $FS$ used in the detection distance calculation leads to more benefits in terms of ERV travel time associated with more non-ERV impacts. With higher factors of safety, non-ERVs are waiting for a lengthier period of time at their final positions.

<table>
<thead>
<tr>
<th>Market Penetration (%)</th>
<th>Percent increase in impact on non-ERVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>6.47%</td>
</tr>
<tr>
<td>40</td>
<td>9.47%</td>
</tr>
<tr>
<td>60</td>
<td>8.28%</td>
</tr>
<tr>
<td>80</td>
<td>7.30%</td>
</tr>
<tr>
<td>100</td>
<td>7.30%</td>
</tr>
</tbody>
</table>

### IX. Conclusion

Sirens and strobe lights have been in use for decades to prioritize ERVs’ movements along streets. Yet, these traditional systems are not efficient when it comes to preventing confusion, as downstream traffic may still be unsure about how to properly react. With the emerging capabilities of ITS technologies, more advanced emergency assistance systems can be developed.

Prior works [1] introduced a system that leverages the connected vehicles technologies to provide link level support to ERVs, an aspect of emergency assistance that has not received enough attention previously. Their system relies on solving ILPs in a sequential manner as the ERV is moving forward along the transportation link. In this paper, the system is extended and combined with other components such as a routing module, to facilitate the movement from an origin to a destination in an urban transportation network with signalized intersections, hence presenting a complete application. After determining the ERV route from origin to destination, a criticality analysis module (CAM), introduced in this paper, identifies the downstream non-ERVs that should receive an instruction message. As the ERV is moving towards its destination, new non-ERVs are detected by CAM. The information of these connected non-ERVs is sent to a module called the sequential ILP optimization (SIOM), which also accommodates the presence of signalized intersections.

To evaluate the benefits and impacts of the proposed system a grid transportation network is modeled in NetLogo, an agent-based modeling environment. Tests were executed for different combinations of congestion and market penetration levels. Results show reduction in ERV travel time (with an average of 9.09% at 100% market penetration), with a minimum recommended market penetration of 40%. More significant reductions are observed in terms of vehicular interactions (with an average of 35.46% and 81.38% for ERV/non-ERV and non-ERV/non-ERV interactions respectively, at 100% market penetration), as the ILP maximizes the gap around the ERV movement and prohibits weaving between non-ERVs. Results confirm that this novel system that is bringing a new aspect of ERV support to light does not only prioritize the ERV’s movement by providing it a safer and clearer passage but also ensures safer movements of downstream vehicles when compared to the currently deployed techniques for ERV assistance.

Further extensions, such as the accommodation of multiple emergencies, will be addressed in future works. To reduce the impact of the ERV’s passage on downstream traffic, driver’s route information will be considered when
available. In addition, techniques for efficient and safe recovery after the ERV’s passage will be developed to assist vehicles resuming their movement. Finally, the implications and countermeasures of imperfect communications will be addressed to prepare this system for short-term deployment.

ACKNOWLEDGEMENTS

This work was supported by NSF through CPS: Synergy: Collaborative Research: Semi-Automated Emergency Response System under Grant 1812524 and Grant 1618979.

REFERENCES


[29] NFPA 1710, standard for the organization and deployment of fire suppression operations, emergency medical operations, and special operations to the public by career fire departments. 2016: National Fire Protection Association.


