Abstract—This paper presents an approach to lane assignment for highway vehicles that increases traffic throughput while ensuring vehicles can exit successfully at their destinations. To enhance traffic safety and increase lane capacities, vehicles can be organized into platoons with the objective of maximizing the travel distance that platoons stay intact and then apply lane assignment to these platoons. The goal of this research is to form a distributed control strategy to select lanes for platoons using inter-vehicle communication. We evaluate the current platoon lane assignment strategy and compare its improvement over average vehicle travel time with the lane assignment for single vehicles reported in our previous work [1]. Simulation results show that while cooperative control for single vehicle lane assignment does lead to decreased vehicle travel times, the implementation of cooperative lane assignment for platooning vehicles leads to an even greater reduction.

I. INTRODUCTION

To define the next-generation of land transportation systems, several transportation projects have been initiated. Examples include the California Partners for Advanced Transit and Highway (PATH) program, the AUTO21 project and other Intelligent Transportation System (ITS) programs. The goals of these programs are quite broad and include increased traffic throughput, less accidents, reduced fuel consumption and better driving experience. However, very little research has been done on developing appropriate algorithms that allow cars to sense and intelligently affect the traffic flows that could result in more efficient use of highways.

Currently, almost all traffic management systems regulate traffic flow by controlling traffic signals or highway ramp meters. In these systems, the traffic is treated as a single mass and the behaviors of individual cars are normally ignored [2][3][4]. Most recently, Goolsby et al. [5] used changeable lane assignment signs at frontage road intersections to adapt to changing traffic conditions at different times of day. When these interchanges experience high turning movement demands, permitted double turns are often used to increase traffic throughput. This approach misses an important component of traffic management: coordination of cars themselves.

This work is funded by AUTO21 Canada.

T. S. Dao is a PhD candidate at the Lab for Autonomous and Intelligent Robotics, Dept. of Mechanical Engineering, University of Waterloo, ON, Canada N2L 3G1 tslao@engmail.uwaterloo.ca

C. M. Clark is an Assistant Professor in the Computer Science Department, California Polytechnic State University, San Luis Obispo, CA, USA 93407 cmclark@calpoly.edu

J. P. Huissoon is a Professor at the Dept. of Mechanical Engineering, University of Waterloo, ON, Canada N2L 3G1 jph@mecheng1.uwaterloo.ca

The work in this paper targets the problem of traffic management using “platoon” concept with the effort to develop scheduling and control techniques to support autonomous driving on urban multi-lane highways with multiple entry (on-ramp) and exit (off-ramp) points. With platooning, several-fold increase in roadway capacity can be achieved with minimal upgrades to infrastructure and relatively little public expense. It is readily apparent that as the speed and density of vehicles increase, the likelihood and likely severity of the crashes will increase. The capabilities of drivers are the principal limitation. Driver errors are responsible for 90% of the crashes that occur today, and the limited ability of drivers to follow other vehicles produces the limitation on lane capacity. The limitation of drivers’ ability to perceive changes in vehicle spacing, relative motion, and acceleration and their limited speed and precision of response ensure that lane capacity cannot generally exceed 2200 vehicles per hour under manual control [6]. In order to increase lane capacity, it is necessary to organize vehicles in platoons where vehicles are at closer average spacing (for the same speed). The platoon mode of operation was conceived as a way of expanding the limitation of capacity and safety that can be achieved by road vehicles. To maximize benefits, it is desirable to form platoons that are reasonably large (five or more vehicles), and it is also desirable to ensure that platoons remain intact for considerable distances [7].

This paper builds from our recent research in the area of lane assignment for single vehicles [1]. Our fundamental assumption is that the cars are equipped with microprocessors, GPS receivers, and wireless communication devices. A following assumption is that any two vehicles within a certain radius of each other can communicate. A car should select lanes not only to improve its own travel time, but the travel time of other cars and the overall traffic flow of the highway system.

Several topics in regard to inter-vehicle communication, platoon assignment, lane finding and lane assignment are explored in this paper: First, the system architecture and the proposed steps to the problem are discussed in Section II. Section III discusses a distributed control strategy for forming platoons and routing platoons to appropriate lanes while satisfying constraint conditions. A cost function to optimize is also proposed. To accurately model traffic flow, a car-following model is proposed in Section IV. Currently, by maximizing the distance that platoons stay intact and applying lane assignment to platoons, the system is able to send cars to appropriate platoons and lanes with an effort to balance lane traffic flows and decrease the vehicles’ travel
time. Simulations to evaluate the algorithm are provided in Section V, followed by some concluding remarks in Section VI.

II. SYSTEM ARCHITECTURE

There are two architectures for implementing lane selecting schemes: centralized and decentralized. The key advantage of decentralized over centralized approach is scalability. It is easier to grow a decentralized system and to add new elements to it.

In this paper, the lane assignment problem for a highway with $n_e$ entry (on-ramp) and $n_d$ exit (off-ramp) points intended for autonomous driving is considered, each lane is characterized by a different nominal driving speed. The highway system used in this work is discretized into segments. For every on-ramp, a new segment is created which contains one or more lanes. The number of lanes can vary from segment to segment but must be a constant along each segment. Lanes are numbered from right to left, with the right-most lane numbered 1. On-ramps and off-ramps are designated as lane 0. Lane exits and lane entrances are assumed to occur on the right side of the highway. It is assumed that a typical vehicle would enter the lanes and proceed gradually to the lane assigned to it.

![Highway system diagram](image)

As illustrated in Fig. 1, the highway is represented by a network. Nodes are assigned to the start of each entry point. The objective is to maximize the flow across the highway. Mathematically speaking, the constraints of the problem are:

1) *Non-negativity*: the number of cars can not be negative.  
2) *Lane capacity*: capacity that determines the maximum rate at which vehicles can enter a lane can not be exceeded.

The proposed implementation of the overall problem can be presented by a closed-loop system and will be broken into four main parts (see Fig. 2): (1) lane positioning (lane occupancy estimation), (2) lane flow estimation, (3) platoon assignment, and (4) lane assignment. Note that each of these individual steps is accomplished through the collaboration of multiple vehicles communicating within ad hoc networks.

Lane positioning is necessary for cars to know where they are on the highway. In an automated highway system, cars are expected to know the lanes they occupy. This serves as a basis for the rest of lane assignment algorithm. A lane positioning algorithm was discussed in detail in [8].

![Control diagram](image)

III. LANE ASSIGNMENT OPTIMIZATION FOR CAR PLATOONS

In this Section, a direction for lane assignment is proposed with an aim to enhance traffic system operation, enhance safety, reduce travel time and improve traffic quality. Both the platoon assignment and lane optimization problems are formulated here as linear programming problems [9].

Due to the limitation of wireless communication range, it is not possible for cars to communicate with all other cars on the highway via ad hoc wireless network. It is necessary to pick out a subset of cars that are able to talk to one another for cooperative planning. Our current approach is to define local clusters of vehicles near the vicinity of road segments. Given a group of cars within the vicinity (defined by $\Delta$ as in Fig. 3) of a road segment, they communicate with each other to create a plan for optimizing traffic throughput in the upcoming road segment (segment $i$). For illustration, call the group of cars within $\Delta$ *group A* (shown in bright color) and the flowing cars that will be within $\Delta$ in the next lane assignment *group B* (shown in dark color). As the first car in group A hits the start of segment $i$, they talk to each other to determine which cars are within $\Delta$. The optimization algorithms are then implemented by individual cars in group A. Once cars in a group have been assigned lanes, the cars will not be reassigned until they form a new group at the vicinity of the next segment. Fig. 3 (bottom) illustrates the lane assignment for cars in group B when the first car in this group hits the start of segment $i$. Details of platoon assignment and lane assignment algorithms are discussed in the next two subsections.
A. Platoon Assignment

This subsection develops a strategy for organizing vehicles into platoons, with the objective of maximizing the distance that platoons stay intact. Fundamentally, this entails grouping vehicles according to their destination.

Let us consider an \( n_p \)-lane highway system as shown in Fig. 1 where \( i, i+1, i+2, \ldots \), indicate the road segment numbers starting at the entry point positions. Let the distance from entry \( i \) to exit \( j \) be \( d_{i,j} \) and the distance that the leading vehicle of platoon \( p \) travel from entry \( i \) to its destination be \( \delta_{p,i} \). The maximum distance a vehicle can travel with platoon \( p \) is \( \min \{ \delta_{p,i}, d_{i,j} \} \) for simplicity. Let \( \chi_{p,j} \) be the number of vehicles that will be sent to platoon \( p \), and \( n_p \) be the number of platoons within \( \Delta \). A candidate cost function to be maximized is:

\[
\Theta = \sum_{p=1}^{n_p} \sum_{j=1}^{n_d} \chi_{p,j} \min \{ \delta_{p,i}, d_{i,j} \} \alpha_{p,j},
\]

where \( \chi_{p,j} \alpha_{p,j} = \chi'_{p,j} \).

The maximization problem can be cast as a linear programming problem to solve for \( \alpha'_{p,j} \)'s with the cost function in Eq. (1) subject to the following constraints:

1) Non-negativity: \( \alpha'_{p,j} \geq 0 \).
2) Maximum platoon size: \( \epsilon_p + \chi_{p,j} \alpha_{p,j} \leq \Upsilon \), where \( \epsilon_p \) is the current size of platoon \( p \), and \( \Upsilon \) is the maximum platoon size (the number of vehicles a platoon can accommodate),
3) Percentages sum to 1: \( \sum_{p=1}^{n_p} \alpha_{p,j} = 1 \).

To summarize, the constrained maximization problem is of the form

\[
\begin{align*}
\max & \ c^T x \\
\text{s.t.} & \ A x \leq b, \ A_{eq} x = b_{eq}, \text{ and } x \geq 0.
\end{align*}
\]

The \( n_p(n_d-i) \times 1 \) vector \( c \) has the form

\[
c = [ \min \{ \delta_{1,d_{1+1}} \} \ldots \min \{ \delta_{n_p,d_{n_p+1}} \} ]^T,
\]

where \( n_p \) components

\[
= \left( \begin{array}{c}
\ldots \\
\min \{ \delta_{1,d_{n_d}} \} \\
\ldots
\end{array} \right),
\]

where \( n_p \) components

\[
\begin{align*}
\text{vector } x & = [ \alpha_{1,i+1} \ldots \alpha_{n_p,i+1} ] \\
\text{matrix } A & = \left[ \begin{array}{ccc}
1 & \cdots & 1 \\
\vdots & \ddots & \vdots \\
1 & \cdots & 1
\end{array} \right],
\end{align*}
\]

and \( A_{eq} \) is a band matrix

\[
A_{eq} = \left[ \begin{array}{ccc}
n_p & 1 & \cdots \\
1 & \ddots & \vdots \\
1 & \cdots & 1
\end{array} \right],
\]

Vectors \( b \) is

\[
b = \left[ \begin{array}{c}
\frac{Y-e_1}{\chi_{i+1}} \\
\vdots \\
\frac{Y-e_{n_d}}{\chi_{n_d}}
\end{array} \right],
\]

and

\[
b_{eq} = \left[ \begin{array}{c}
1 \\
\vdots \\
1
\end{array} \right],
\]

B. Lane Assignment for Platoons

To assign lanes to platoons, we consider vehicle platoons as single vehicles and apply the lane assignment algorithm to platoons as done for single vehicles in [1].

The lane assignment problem can be summarized as follows: Let the nominal velocity on lane \( c \) be \( v \). The estimated time that takes one platoon currently in lane \( l_c \) to switch to lane \( l_c \), then travel from \( i \) to exit \( j \) is \( \frac{d_{i,j}}{v} + T_{i,j} \). Note that \( T_{i,j} = \tau|l_l - l_c| + \tau l \) is the time delay caused by lane change maneuvers considering that the maneuver cost (time penalty) for a platoon to move from lane \( l_c \) to lane \( l_l \) is \( \tau|l_l - l_c| \) and from lane \( l_l \) to an off-ramp (exiting) is \( \tau l \). Here, \( \tau \) is the predefined time constant (off-ramps are considered as lane 0).

The total time of all cars in lane \( l_c \) starting at segment \( i \) is \( \sum_{j=1}^{n_d} \rho_{l_c,l_l}^{l_c,j} (\frac{d_{j}}{v} + T_{i,j}) \), where \( \rho_{l_c,l_l}^{l_c,j} \) is the number of platoons currently in lane \( l_c \) that will be traveling from the start of segment \( i \) to exit \( j \) in lane \( l_l \), and \( n_d \) is the number of exit points. Thus the cost function to be minimized is:

\[
\sum_{l_c=0}^{\mu} \sum_{j=1}^{n_d} \sum_{j=1}^{n_d} \rho_{l_c,l_l}^{l_c,j} (\frac{d_{j}}{v} + T_{i,j}).
\]
The problem is to assign lanes to all platoons within ∆ at the beginning of each segment. Let \( \beta_{lc,li,j} \) be the percentage of platoons (within \( \Delta \)) traveling from \( i \) to \( j \) that will be sent from lane \( l_c \) to lane \( l_i \). The factor \( \beta_{lc,li,j} \) relates the number of platoons \( \rho_{lc,li,j} \) by \( \rho_{lc,li,j} = \beta_{lc,li,j} \rho_{li,j} \). The cost function for segment \( i \) becomes

\[
\Phi_i = \sum_{l_i=0}^{n_i} \sum_{l_j=1}^{n_j} \sum_{c=1}^{n_c} \rho_{lc,li,j} \left( \frac{d_{li,j}}{v} + T_{li,j} \right) \beta_{lc,li,j}.
\] (2)

The constraints to the problem are:
1) Non-negativity: \( \beta_{lc,li,j} \geq 0 \),
2) Lane capacities: capacities of lanes should not be exceeded,
3) Percentages sum to 1: \( \sum_{j=1}^{n_j} \sum_{c=1}^{n_c} \beta_{lc,li,j} = 1 \).

More details about lane assignment algorithm can be found in [1].

IV. SPEED-FLOW-DENSITY RELATIONSHIP

The total travel time can be better estimated by taking the speed-flow-density relationship into account. To formulate Eq. (2), the vehicle speeds are assumed to be some nominal lane velocities. However, in an actual highway, the vehicle speed is dependent on the traffic density (and hence the flow rate).

In this work, Van Aerde’s model [10] is adopted and implemented in the traffic simulator to model the speed-flow-density relationship. This integration model that uses a steady-state car-following model proposed by Van Aerde combines of the Pipes and Greenshields models [11] into a single-regime model. The model, which requires three input parameters, can be calibrated using field loop detector data. The efforts for calibrating Van Aerde’s model were described in detail in [10]. The Van Aerde single-regime model overcomes the shortcomings of Greenshields and Pipes models which are often inconsistent with field data from a variety of highways (see [11]).

Van Aerde’s model can be described by a series of expressions as follows

\[
\gamma = \frac{1}{c_1 + \frac{c_2}{v_f} + c_3v},
\] (3a)

\[
m = \frac{2v_c - v_f}{(v_f - v_c)^2},
\] (3b)

\[
c_2 = \frac{1}{\gamma_f(m + \frac{1}{v_f})},
\] (3c)

\[
c_1 = mc_2,
\] (3d)

\[
c_3 = \frac{-c_1 + \frac{v_c}{\gamma_f} - c_2}{v_c},
\] (3e)

where:

\( \gamma = \) traffic density (veh/km) or the inverse of the vehicle headway (km/veh),

\( v = \) vehicle speed (km/h),

\( v_f = \) free-speed (km/h),

\( v_c = \) speed at capacity (km/h),

\( \gamma_f = \) jam density (veh/km),

\( c_1 = \) fixed distance headway constant (km),

\( c_2 = \) first variable headway constant (km²/h),

\( c_3 = \) second variable headway constant (h⁻¹).

In practice, the calibration of the car-following model requires the estimation of three parameters: \( v_f, v_c, \) and \( \gamma_f \).

The vehicle speed can be inferred from traffic density \( \gamma \) and flow rate \( F \) using the fundamental speed-flow-density relationship

\[
v = \frac{F}{\gamma}.
\] (4)

From Eq. (3a), the traffic density \( \gamma \) can be expressed as a function of flow by replacing \( v \) with Eq. (4). This gives the quadratic equation

\[
\gamma^2(c_1v_f + c_2) + \gamma(-c_1F + c_3Fv_f - v_f) + (F - c_3F^2) = 0.
\] (5)

The estimate vehicle speed \( v_f \) in Eq. (2) (represented by \( v \) in Eqs. (3)-(4)) is a function of flow and can be obtained by solving Eq. (5) and substituting the resulting \( \gamma \) into Eq. (4). In summary, the procedure is to measure the flow rate \( F \) for each lane, then solve for \( v \) and use it in the cost function in Eq. (2) as the nominal lane velocities.

V. SIMULATION RESULTS ANALYSIS

To implement simulations, the VISSIM software package [12] was used. VISSIM is a microscopic, time step and behavior based simulator developed to analyze the full range of functionally classified roadways. It is capable of modeling traffic with various control measures in a 3D environment. VISSIM lets us communicate and control the behaviors of vehicles through a dynamic link library (DLL) file compiled from C/C++ code. Vehicle parameters from the external driver model DLL output function are stored within member variables of a designated vehicle class object. A detailed description about the software architecture used in this research is provided in [13].

Based on the platoon assignment and lane assignment strategies described in Section III, it is possible to assign appropriate lanes to cars. Up to this point, our system assigns cars to platoons then assign platoons to lanes once the vehicles come within the vicinity of the next road segment. The average travel time of all the vehicles on the highway is recorded to evaluate the effectiveness of the algorithm.

The four-lane highway used in the simulation has three entry and three exit points. The maximum capacity for each of the four freeway lanes is 2200 veh/h at an average of 1.6 seconds between vehicles. The highway starts with zero traffic and the vehicles are generated randomly by VISSIM. The speed-flow relationships in all four lanes of
the highway follow Van Aerde’s model which are similar to the calibration results from a real freeway described in [10]. The free-speeds for lane 1 through lane 4 are 80 km/h, 90 km/h, 100 km/h and 120 km/h, respectively. The speeds at capacity \( v_c \) in the four lanes are 81% of the free-speeds. To ensure the consistency between simulations with and without lane assignment, the simulation parameters such as speed-flow-density model (Van Aerde’s model), positions of lane change decision, minimum lane acceptance gaps, etc., are the same for all simulations. The only thing making the two strategies different is that vehicles’ lane changing behavior is completely controlled by VISSIM in simulations without lane assignment. That is, a driver has a desire to change lane if he has to drive slower than his desired speed due to a slow leading vehicle or in case of an upcoming exit with a special deceleration lane. First, the driver checks whether he can improve his present situation by changing lanes. Then he checks whether he can change without generating a dangerous situation [14].

When vehicles are joined into a closely spaced platoon, the gap between vehicles in platoons are reduced to \( h = 10m \). Let the time headway between Platoons be \( t_h = 5s \), the capacity of lane \( l \) is increased to \( C_{max} = \frac{3600t_h}{h} \). For example, the lane capacity can be increased up to 7200 vehicles per lane per hour when the maximum platoon size of 25 vehicles is used.

The effect of platooning can be seen in Fig. 4. This figure shows the average travel time for 1000 vehicles on the highway with different total input volumes. For comparison, results from different simulations in which no lane assignment strategy is applied (i.e., vehicles can freely choose lanes). Also, the lane assignment algorithm is applied for single vehicles [1]. Finally, the platoon assignment and lane assignment are applied with the maximum platoon size of 25 vehicles and the results are included in this figure. The total volumes at the inputs in the three simulations are 2000 veh/h, 4000 veh/h and 6000 veh/h, respectively. The recorded total number of lane changes varies between 1800~1900 for no lane assignment, 3100~3150 for single vehicle lane assignment, and 3050~3700 for platoon lane assignment while the input volume takes on values from 2000 to 6000 veh/h. It can be seen that platooning gives greater traffic flow at the cost of increased lane changes. The single vehicle lane assignment strategy also has this property, but to a lesser extent than platooning.

The first case (Fig. 4 (top)) indicates that there’s not much improvement compared to lane assignment for single vehicles in considerably light traffic conditions. However, in heavy traffic (Fig. 4 (middle and bottom)), vehicle travel time can be further reduced when both platooning and lane assignment are implemented. This can be explained by the fact that in light traffic, fast lanes can accommodate all vehicles no matter what lane assignment strategy is applied while in highly congested highway, the number of vehicles that can enter fast lanes is limited by lane capacities (which are low when platooning is not used).

The dependency of vehicle travel time on maximum platoon size (and hence lane capacities) is illustrated in Fig. 5. This figure plots average travel time for 1000 vehicles with the maximum platoon size varying from 5 to 30 vehicles and total input volume at all entrance ramps being 10,000 veh/h. It can be seen that as the maximum platoon size increases, the average vehicle traffic time is reduced. This is obvious since the capacities of fast lanes are now increased and they can allow more vehicles to enter.

The maximum platoon size, however, will be limited by the risk of collisions since an accident involving any vehicle in a platoon will also impact its following vehicles. In order to promote public acceptance, the danger from driving in a platoon should be substantially less than the equivalent risk of driving in a passenger car today. Cooperative intersection collision avoidance systems [15][16], one-way streets for platoons, etc., are examples of ways to reduce this risk.

Fig. 4. Improvement over average travel time for 1000 cars with the input volume being 2000 veh/h (top), 4000 veh/h (middle), and 6000 veh/h (bottom).

Fig. 5. Average travel time vs platoon size.
VI. CONCLUSION

An algorithm for optimization of platoon assignment and lane assignment was developed with the effort to increase traffic throughput. We have proposed linear models for both optimization problems and presented preliminary results on the use of a linear programming algorithm in the solution of the problem. It has been shown that with platooning, several-fold increase in lane capacities can be achieved when vehicles are clustered together in groups with very short spacings between vehicles within platoons and long spacings between platoons.

The main outcome of this work is to improve both traffic throughput and safety by the use of automation since electronic sensors, computers, and actuators can provide faster and more precise responses than human drivers.

Simulation results suggest that by maximizing the travel distance that platoons stay intact and then applying appropriate lane assignment algorithm to the platoons, vehicle travel times can be reduced even greater compared to cooperate control for single vehicle lane assignment.

VII. ACKNOWLEDGMENTS

Financial support for this research has been provided by AUTO21 Canada.

REFERENCES