Chapter

Models and Methods for Intelligent Highway Routing of Human-Driven and Connected-and-Automated Vehicles

Fatemeh Alimardani, Nilesh Suriyarachchi,
Faizan M. Tariq and John S. Baras

Abstract

Connected and automated vehicles (CAVs) have seen a rapid surge in interest over the past few years. A lot of focus is being placed on improving the efficiency and robustness of transportation systems by leveraging the sensors and capabilities of CAVs. However, the integration of CAVs into existing traffic infrastructure would give rise to certain issues that must be addressed before the CAVs can be seen ubiquitously on public roads. Since the highway networks are considered permanent investments that are expensive to build and maintain, the priority is to improve the efficiency of the current traffic system. This chapter explores the integration of two of the most common traffic management strategies, namely, ramp metering (RM) and route guidance (RG), into existing highway networks with human-driven vehicles (HDVs). The introduction of CAVs to public roads will engender issues pertaining to safe interactions between CAVs and HDVs. The later part of the chapter addresses the specific problems of improving highway on-ramp merging efficiency by optimally coordinating CAVs. The chapter concludes by presenting a scenario that requires an explicit consideration of interactions between HDVs and CAVs.

Keywords: ramp metering, route guidance, merging behavior, overtaking behavior, human-driven vehicles, connected-and-automated vehicles

1. Introduction

The ultimate goal of automating the driving process is to improve safety by reducing accidents caused by human errors. If all vehicles in a network are human–driven, the efficiency of traffic networks can be improved by the control of traffic signal lights and the routes that drivers can choose. Studying the literature on freeway traffic control for HDVs demonstrate that the integration of traffic control strategies such as ramp metering (RM) and route guidance (RG) improve the network performance in regards to travel time, travel distance, throughput and emissions.

Moreover, it seems unrealistic that all the HDVs will suddenly be replaced by the AVs in the near future. Rather, what seems more plausible is that the AVs will be introduced onto the roads in the presence of the HDVs. Therefore, there is a need to consider cases where it becomes necessary to model the interactions between AVs and HDVs. Delays caused at on-ramps and off-ramps are some of the major contributors to overall system efficiency degradation. In addition to the increase of congestion in the merge lane and outer freeway lanes, merging lanes can have an overflow effect which causes the entire freeway to become congested. However, with the advent of CAVs, a lot more information has been made available for improving this overall process.

Moving on to mixed–autonomy highway networks, as a specific example of the interaction between HDVs and CAVs, the overtaking behavior performed by a CAV is chosen as the target driving behavior for the last section of this chapter. The reason for this choice is that it is one of the more challenging driving behaviors when compared to car following and lane changing as it encompasses the combination of these behaviors.

This chapter is organized as follows: Section 2 reviews the integration of RM and RG, that have shown significant improvements on different control measures for highway networks with HDVs. Section 3 addresses the specific problem of improving freeway on-ramp merging efficiency by optimally coordinating CAVs. Finally, Section 4 explores the overtaking behavior accomplished by a CAV in the presence of HDVs.

2. Integration of ramp metering and route guidance for HDVs

This section will focus on providing a review on the combined RM and RG control as two of the most common traffic control management techniques. To do so, first, a review on traffic flow models will be provided and then, the most common RM and RG strategies will be explained, respectively. At the end of this section, a review on the studies with the focus on the integration of RM and RG will be presented.

2.1 Traffic flow models

Traffic flow models can be categorized into first order and second order models. The most frequently used models are first order models, such as Lighthill-Whitham-Richards (LWR) model [1], which is a continuous model, and the cell-transmission model (CTM) [2], which is a discretized version of the LWR model. The second-order traffic flow models, besides considering the dynamics of the traffic density, introduce a dynamic equation for the mean velocity. The most famous second order model is the Modèle d'Écoulement de Trafic sur Autoroute NETworks (METANET) model [3, 4]. In this section, a review on the CTM and METANET model, as the two most used discrete traffic flow models in the literature, will be provided. The notations adopted in this section are adopted from [5]. **Tables 1** and **2** describe the model variables and parameters of these two models with their symbols, definitions, and units.

2.1.1 The cell transmission model (CTM)

The CTM was first developed by Daganzo [2] in 1992 and then, through out the following years, many other extensions of it were developed. The following version is the original version of the CTM with some minor modifications from [6] and the notations are borrowed from [5]. The CTM is characterised by the following equations:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L}\left(\Phi_i^+(k) - \Phi_i^-(k)\right) \tag{1}$$

$$\Phi_i^+(k) = \phi_i(k) + r_i(k) \tag{2}$$

$$\Phi_i^-(k) = \phi_{i+1}(k) + s_i(k) \tag{3}$$

$$s_i(k) = \frac{\beta_i(k)}{1-\beta_i(k)}\phi_{i+1}(k) \tag{4}$$

The dynamic equation of the on-ramp queue length is:

$$l_i(k+1) = l_i(k) + T(d_i(k) - r_i(k)). \tag{5}$$

The mainline flows and on-ramp flows are:

$$\phi_i(k) = \min\left\{(1-\beta_{i-1}(k))v_{i-1}(\rho_{i-1}(k) + r_{i-1}(k)), w_i\left(\rho_i^{max} - \rho_i(k) - r_i(k)\right), q_i^{max}\right\} \tag{6}$$

Symbol	**Description**	**Unit/Range**
T	Sampling time	$[h]$
N	Number of cells	int
i	Cell index	$i = \{1, \dots, N\}$
K	Time horizon	int
k	Time index	$k = \{0, \dots, K-1\}$
L	Length of each cell	$[km]$
v_i	Free-flow speed	$[km/h]$
ω_i	Congestion wave speed	$[km/h]$
q_i^{max}	Cell capacity (Maximum flow rate)	$[veh/h]$
ρ_i^{max}	Jam density	$[veh/km]$
ρ_i^{cr}	Critical density	$[veh/km]$
l_i^{max}	Maximum on-ramp queue length	$[veh]$
$r_i^{C,max}$	Maximum ramp metering rate	$[veh/h]$
$\rho_i(k)$	Traffic density	$[veh/km]$
$\Theta_i^+(k)$	Total flow entering cell i	$[veh/h]$
$\Theta_i^-(k)$	Total flow exiting cell i	$[veh/h]$
$\phi_i(k)$	Mainstream flow entering cell i from cell $i-1$	$[veh/h]$
$r_i(k)$	Flow entering cell i from its on-ramp	$[veh/h]$
$s_i(k)$	Flow exiting cell i through its off-ramp	$[veh/h]$
$\beta_i(k)$	Split ratio	$\in [0,1]$
$l_i(k)$	Queue length in the on-ramp	$[veh]$
$d_i(k)$	Flow accessing the on-ramp	$[veh/h]$
$r_i^C(k)$	Ramp metering control variable	$[veh/h]$

Table 1.
CTM model variables and parameters of cell i during interval [kT,(k+1)T).

Symbol	Description	Unit/Range
T	Sampling time	$[h]$
K	Time horizon	int
k	Time index	$k = \{0, \dots, K-1\}$
M	Number of mainline links	int
m	Mainline link index	$m = \{1, \dots, M\}$
N_m	Number of sections of mainline link m	int
i	Section index	$i = \{1, \dots, N_m\}$
O	Number of origin links	int
o	Origin link index	$i = \{1, \dots, O\}$
L_m	Length of each mainline link m	$[km]$
λ_m	Lane numbers of each mainline link m	int
O_n	Set of exiting mainline links from node n	-
I_n	Set of entering mainline links to node n	-
$\bar{I}_n$	Set of entering origin links to node n	-
J_m	Set of destinations reachable from mainline link m	-
$\bar{J}_o$	Set of destinations reachable from origin link o	-
$\bar{\bar{J}}_n$	Set of destinations reachable from node n	-
$v^f_{m,i}$	Free-flow speed in section i of link m	$[km/h]$
ρ^{cr}_m	Critical density	$[veh/km]$
ρ^{max}_m	Jam density	$[veh/km]$
q^{max}_o	Capacity of origin link o	$[veh/h]$
τ	Model parameter	-
η	Model parameter	-
χ	Model parameter	-
ϕ	Model parameter	-
a_m	Model parameter	-
δ_{on}	Model parameter	-
$\rho_{m,i,j}(k)$	Partial density	$[veh/km]$
$\rho_{m,i}(k)$	Total density	$[veh/km]$
$v_{m,i}(k)$	Mean traffic speed	$[km/h]$
$q_{m,i}(k)$	Traffic flow leaving section i of link m	$[veh/h]$
$\gamma_{m,i,j}(k)$	Composition rate	$\in [0, 1]$
$d_{o,j}(k)$	Partial origin demand at origin link o	$[veh/h]$
$d_o(k)$	Total origin demand at origin link o	$[veh/h]$
$l_{o,j}(k)$	Partial queue length at origin link o	$[veh]$
$l_o(k)$	Total queue length at origin link o	$[veh]$
$\gamma_{o,j}(k)$	Composition rate	$\in [0, 1]$
$\theta_{o,j}(k)$	Portion of demand originating in origin link o	$\in [0, 1]$

$q_o(k)$	Total traffic volume leaving origin link o	$[veh/h]$
$r_o^C(k)$	Ramp metering control variable	$[veh/h]$
$Q_{n,j}(k)$	Flow entering node n	$[veh/h]$
$\beta_{m,n,j}(k)$	Split ratio	$\in [0,1]$

Table 2.
METANET model variables and parameters of mainline link m, section i, node n, origin link o, destination j during interval [kT,(k+1)T).

$$r_i(k) = \begin{cases} \min\left\{l_i(k) + d_i(k), \rho_i^{max} - \rho_i(k)\right\} & \text{Uncontrolled On} - \text{Ramps} \\ \min\left\{l_i(k) + d_i(k), \rho_i^{max} - \rho_i(k), r_i^{C,max}\right\} & \text{Controlled On} - \text{Ramps} \end{cases} \tag{7}$$

Metering rate variables $r_i^C(k)$ come from the RM control law which will be mentioned in detail in Section 2.2. All variables are bounded between zero and their maximum possible value.

Many extensions of the original CTM have been proposed in the literature in the last two decades. The CTM in a mixed-integer linear form [7], the CTM including capacity drop phenomena [8, 9], the CTM for a freeway network [10], the asymmetric CTM [6], the link-node CTM [11], and the variable-length CTM [12] are some of these extended versions. Although these models have been proposed in different years and are suitable for different networks and applications, the original CTM [2] is the underlying model in all of them and it proves how powerful the original CTM is.

2.1.2 The METANET model

The METANET model presented here is an improved version [4] of the original that was first presented in [3]. However, the notation has been adopted from [5] in order to agree with the other notations of this section.

Freeway Links

$$\rho_{m,i,j}(k+1) = \rho_{m,i,j}(k) + \frac{T}{L_m \lambda_m}\left[\gamma_{m,i-1,j}(k) q_{m,i-1}(k) - \gamma_{m,i,j}(k) q_{m,i}(k)\right] \tag{8}$$

$$\rho_{m,i}(k) = \sum_{j \in J_m} \rho_{m,i,j}(k) \tag{9}$$

$$\gamma_{m,i,j}(k) = \frac{\rho_{m,i,j}(k)}{\rho_{m,i}(k)} \tag{10}$$

$$\nu_{m,i}(k+1) = \nu_{m,i}(k) + \frac{T}{\tau}\left[V\left(\rho_{m,i}(k)\right) - \nu_{m,i}(k)\right] + \frac{T}{L_m}\nu_{m,i}(k)\left[\nu_{m,i-1}(k) - \nu_{m,i}(k)\right] - \frac{\nu' T\left[\rho_{m,i+1}(k) - \rho_{m,i}(k)\right]}{\tau L_m\left[\rho_{m,i}(k) + \chi\right]} \tag{11}$$

$$q_{m,i}(k) = \rho_{m,i}(k)\nu_{m,i}(k)\lambda_m \tag{12}$$

$$V\left(\rho_{m,i}(k)\right) = \nu_m^f \exp\left[-\frac{1}{a_m}\left(\frac{\rho_{m,i}(k)}{\rho_m^{cr}}\right)^{a_m}\right] \tag{13}$$

The speed reduction caused by merging phenomena near on-ramps (possible additional term to Eq. (11)):

$$-\delta_{on} T \frac{\nu_{m,1}(k) q_o(k)}{L_m \lambda_m \left[\rho_{m,1}(k) + \chi\right]} \tag{14}$$

The speed reduction due to weaving phenomena in case of lane reductions in the mainstream (possible additional term to Eq. (11)):

$$-\phi T \Delta\lambda \frac{\nu_{m,N_m}(k)^2 \rho_{m,N_m}(k)}{L_m \lambda_m \rho_m^{cr}} \tag{15}$$

The virtual downstream density at the end of the link (for node n at the end of link m with more than one outgoing link):

$$\rho_{m,N_m+1}(k) = \frac{\sum_{\mu \in O_n} \rho_{\mu,1}(k)^2}{\sum_{\mu \in O_n} \rho_{\mu,1}(k)} \tag{16}$$

The virtual upstream speed at the beginning of the link (for node n at the beginning of link m with more than one entering freeway link):

$$\nu_{m,0}(k) = \frac{\sum_{\mu \in I_n} \nu_{\mu,N_\mu}(k) q_{\mu,N_\mu}(k)}{\sum_{\mu \in I_n} q_{\mu,N_\mu}(k)} \tag{17}$$

Origin links

$$l_{o,j}(k+1) = l_{o,j}(k) + T\left[d_{o,j}(k) - \gamma_{o,j}(k) q_o(k)\right] \tag{18}$$

$$l_o(k) = \sum_{j \in \bar{J}_o} l_{o,j}(k) \tag{19}$$

$$\gamma_{o,j}(k) = \frac{l_{o,j}(k)}{l_o(k)} \tag{20}$$

$$d_{o,j}(k) = \theta_{o,j}(k) d_o(k) \tag{21}$$

For uncontrolled on-ramps:

$$q_o(k) = \min\left\{d_o(k) + \frac{l_o(k)}{T}, q_o^{max}, q_o^{max} \frac{\rho_m^{max} - \rho_{m,1}(k)}{\rho_m^{max} - \rho_m^{cr}}\right\} \tag{22}$$

For controlled on-ramps:

$$q_o(k) = \min\left\{d_o(k) + \frac{l_o(k)}{T}, q_o^{max}, r_o^C(k), q_o^{max} \frac{\rho_m^{max} - \rho_{m,1}(k)}{\rho_m^{max} - \rho_m^{cr}}\right\} \tag{23}$$

where $r_o^C(k)$ come from the RM control law which will be mentioned in detail in Section 2.2.

Nodes

$$Q_{n,j}(k) = \sum_{\mu \in I_n} q_{\mu,N_\mu}(k) \gamma_{\mu,N_{\mu,j}}(k) + \sum_{o \in \bar{I}_n} q_o(k) \gamma_{o,j}(k) \tag{24}$$

$$q_{m,0}(k) = \sum_{j \in J_m} \beta_{m,n,j}(k) Q_{n,j}(k) \tag{25}$$

$$\gamma_{m,0,j}(k) = \frac{\beta_{m,n,j}(k) Q_{n,j}(k)}{q_{m,0}(k)} \tag{26}$$

In presence of RG control, the splitting rates become the control variables and are calculated based on the RG control law. It will be mentioned in detail in Section 2.3.

Few research studies have developed different versions of the METANET model due to the complexity and non-linearity of the second-order models. However, the extensions of the METANET for a freeway network [13], and the multi-class METANET both for a freeway stretch [14] and for a freeway network [15] are examples of the extensions of this second-order traffic model developed in the recent years.

Both the first-order and second-order models are capable of developing the evolution of traffic flow in both urban and non-urban network. However, to highlight their difference, it is necessary to emphasize that first order models focus on the evolution of the density while the second-order traffic flow models, besides considering the dynamics of the traffic density, explicitly introduce a dynamic equation for the mean speed. Second order models have the distinct advantage over first order models that they can reproduce the capacity drop, which is the observed difference between the freeway capacity and the queue discharge rate. First order models, because they do not capture this phenomenon, are incapable of exploiting the benefits of increasing bottleneck flow. They can only reduce travel time by increasing off-ramp flow. The obvious disadvantage to second order models is that they lead to more complex optimization problems.

The focus of the rest of this section will be on ramp metering and route guidance control schemes as two of the most famous traffic management techniques.

2.2 Ramp metering

Ramp metering is achieved by placing traffic signals at on-ramps to control the flow rate at which vehicles enter the freeway. The ramp metering controller computes the metering rate to be applied. Ramp metering has various goals [16]: to improve or remove congestion, to alleviate freeway flow, traffic safety and air quality, to reduce total travel time and the number of peak-period accidents, to regulate the input demand of the freeway system so that a truly operationally balanced corridor system is achieved. Although the ramp metering provides many advantages, at the same time, it can have disadvantages too. The following are two of the most plausible ones [5]: (1) drivers may use parallel routes to avoid ramp meters which may lead to increased travel time and distance, (2) it can shift the traffic congestion from one location to another.

Ramp metering control strategies can be classified in the following categories [16]: (1) local system where the control is applied to a single on-ramp, (2) coordinated system where the control is applied to a group of on-ramps, considering the traffic conditions of the whole network, (3) integrated system where a combination of ramp metering, signal timing, and route guidance is applied as the control system. Also, from another point of view, there are two types of RM control schemes [16]: (1) pre-timed or isolated where metering rates are fixed and predefined, (2) traffic-responsive control where real time freeway measurements are used to determine the control variables.

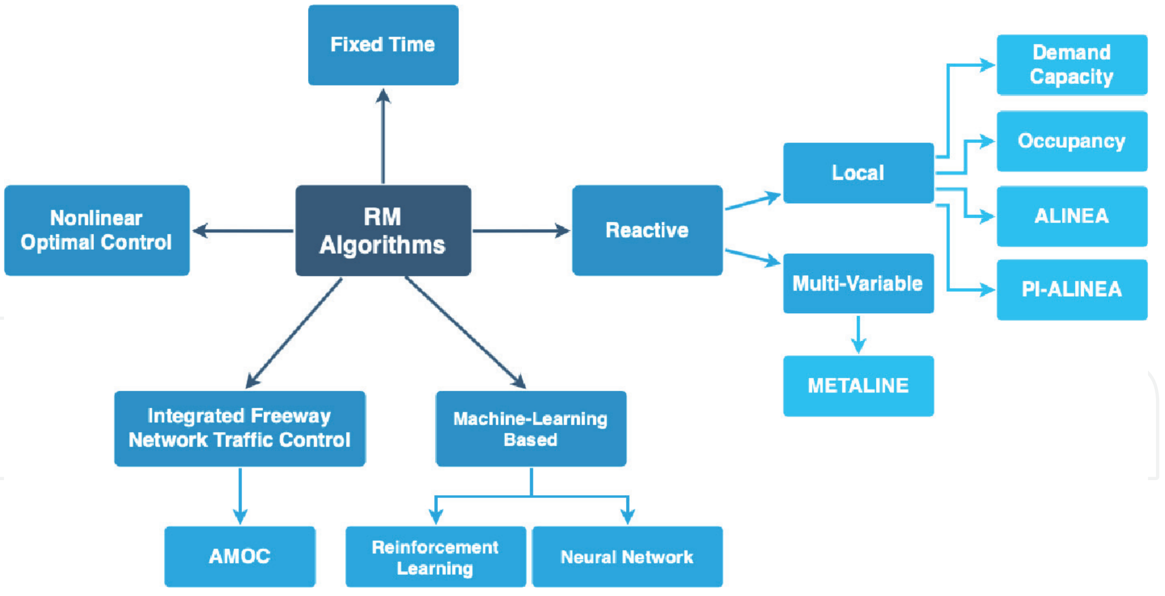

Figure 1.
Ramp metering algorithms classification.

A classification of ramp metering algorithms based on a study by Papageorgiou and Kotsialos [17] is presented in **Figure 1**. Fixed time metering is the simplest strategy which is usually adjusted based on historical data and applied during particular times of day. Reactive ramp metering techniques are based on real time traffic metrics. Local ramp metering uses traffic measures collected form the ramp vicinity. Demand-capacity, and occupancy-based strategies allow as much traffic inflow as possible to reach the freeway capacity. ALINEA and PI-AlINEA offer a more complex and more responsive strategy that, unlike capacity and occupancy strategies, generates smoother responses towards changes in metrics. Multi-variable regulator strategies perform the same as local strategies, but more comprehensively and independently on a set of ramps and usually outperform local strategies. METALINE can be viewed as a more general and extended form of ALINEA. Nonlinear optimal control strategy considers local traffic parameters and metrics as well as nonlinear traffic flow dynamics, incidents, and demand predictions in a freeway network and outputs a consistent control strategy. Knowledge-based control systems are developed based on historical data and human expertise. Integrated freeway network traffic control is a more general approach to nonlinear control that extends application of optimal control strategies to all forms of freeway traffic control. In case of knowledge based systems, inability to learn and adapt to temporal evolution of the system being controlled can be an issue, so knowledge based systems need to be periodically updated to remain efficient. Artificial intelligence and machine learning approaches like reinforcement learning (RL) and artificial neural networks (ANN) are new techniques being implemented recently for RM control [18].

Two of the most common ramp metering strategies are described in the following. Here, the flow that can enter section i of a freeway from the on-ramp during time interval $[kT, (k+1)T)$ is shown by $r_i^C(k)$ where k is the time index and T is the sampling time.

2.2.1 ALINEA

ALINEA [19] is an I-type controller in which the metering rate is given by

$$r_i^C(k) = r_i^C(k-1) - K_R\left[\rho_i^* - \rho_i^{down}(k)\right] \tag{27}$$

where $\rho_i^{down}(k)$ is the density measured downstream the on-ramp, ρ_i^* is a set-point value for the downstream density, and K_R is the integral gain. Note that, in case the main objective of the traffic controller is to reduce congestion and to maximise the throughput, a good choice for the set-point is $\rho_i^* = \rho_i^{cr}$.

2.2.2 PI-ALINEA

A very famous extension of ALINEA is the PI-ALINEA, in which a proportional term is added to result in a PI regulator. The metering rate is given by

$$r_i^{\mathrm{C}}(k) = r_i^{\mathrm{C}}(k-1) - K_P\left[\rho_i^{down}(k) - \rho_i^{down}(k-1)\right] + K_R\left[\rho_i^* - \rho_i^{down}(k)\right] \tag{28}$$

where K_P is another regulator parameter.

Based on the stability analysis of the closed-loop ramp metering system provided in [20], it can be stated that PI-ALINEA is able to show a better performance than ALINEA.

2.3 Route guidance

Route guidance (RG) is an efficient technique to distribute the traffic demand over the network, by providing information about alternative paths to drivers. The variable message signs (VMSs) are one of the main actuators which can provide route information to drivers in the RG control scheme. In the RG control, the concepts of equilibrium play an important role. Wardrop has offered the two following principles [21]: (1) The *system optimum (SO)* is achieved when the vehicles are guided such that the total costs of all drivers (typically the TTS) is minimized, (2) The traffic network is in *user equilibrium (UE)* when the costs on each utilized alternative route is equal and minimal, and on routes that are not utilized, the cost is higher that on the utilized routes.

If the goal of a control strategy is defined as the travel time, it is typically defined as the *predicted travel time* or as the *instantaneous travel time*. The predicted travel time is the time that the driver will experience when he drives along the given route, while the instantaneous travel time is the travel time determined based on the current speeds on the route. In a dynamic setting, the instantaneous travel time may be different from the predicted travel time [5].

In route guidance control strategies, the control variable is the splitting rate at a given node. Considering the simple case of only two alternative paths [22], originating from node n, let us denote with m and m' the two links exiting node n, corresponding respectively to the primary and secondary path. The primary path is the one characterised by the shortest travel time, in case of regular traffic conditions. In particular, the control variable is the splitting rate $\beta_{m,n,j}^{\mathrm{C}} \in [0, 1]$, representing the portion of flow present in node n at time instant kT which should choose link m to reach destination j. The other control variable is $\beta_{m',n,j}^{\mathrm{C}}$, referred to link m', where it is easily computed: $\beta_{m',n,j}^{\mathrm{C}} = 1 - \beta_{m,n,j}^{\mathrm{C}}$. The following feedback regulators of P-type or PI-type are the most used strategies for route guidance systems in the literature [5, 23]. According to a proportional control law, the portion of flow present in node n at time instant kT which should choose link m to reach destination j is computed as

$$\beta_{m,n,j}^{\mathrm{C}}(k) = \beta_{m,n,j}^{N}(k) + K_P \Delta\tau_{n,j}(k) \tag{29}$$

where $\beta^{N}_{m,n,j}(k)$ is the nominal splitting rate, K_P is a gain, $\Delta\tau_{n,j}(k)$ is the instantaneous travel time difference between the secondary and primary direction from n to j. In proportional-integral regulators, the splitting rate is

$$\beta^{C}_{m,n,j}(k) = \beta^{C}_{m,n,j}(k-1) + K_P\left[\Delta\tau_{n,j}(k) - \Delta\tau_{n,j}(k-1)\right] + K_I\Delta\tau_{n,j}(k) \tag{30}$$

where K_P and K_I are other controller gains.

Another possible class of RG strategies is *iterative strategies*, where the splitting rate is computed by iteratively running different simulations in real time with different RG, in order to achieve conditions of either user equilibrium or system optimum [24, 25]. Iterative strategies are very beneficial, however, their high computational effort is a major drawback for this category of RG control techniques.

It is also interesting to describe how drivers react to travel time information and how they adapt their route choice. A well-known behavior model used for this purpose is the logit model [26], which is used to model all kinds of consumer behavior based on the cost of several alternatives. The lower the cost of an alternative, the more consumers will choose that alternative. In the case of traffic management, consumers are the drivers, and the cost is the comfort, safety, or travel time of the alternative routes to reach the desired destination. The logit model calculates the probability that a driver chooses one of more alternatives based on the difference in travel time between the alternatives. Assume that we have two possible choices m_1 and m_2 at node n to get to destination j. For the calculation of the split rates out of the travel time difference between two alternatives, the logit model results in:

$$\beta_{m,n,j}(k) = \frac{\exp\left(\sigma\theta_{n,m,j}(k)\right)}{\exp\left(\sigma\theta_{n,m_1,j}(k)\right) + \exp\left(\sigma\theta_{n,m_2,j}(k)\right)} \tag{31}$$

for $m = m_1$ or $m = m_2$, where $\theta_{n,m_1,j}(k)$ is the travel time shown on the DRIP at node n to travel to destination j via link m. The parameter σ describes how drivers react on a travel time difference between two alternatives. The higher σ, the less travel time difference is needed to convince drivers to choose the fastest alternative route.

2.4 Integration of ramp metering and route guidance

The RM and RG controllers are both feedback and predictive controllers as they apply not only on the real-time measurements of the system to calculate the control actions, but they also use information about the prediction of the system evolution. The combination of RM and RG controllers has shown promising results in network performance from the point of view of different performance measures. The focus of this section is on providing a review on the related studies on the combination of these two main traffic management techniques.

In 1990, a study performed by Iida et al. [27] considered the development of an improved on-ramp traffic control technique of urban expressway. They extended the conventional LP control method to consider the multiple paths between on-ramps and off-ramps of the test case network and also the route choice behavior of drivers. They assumed that in the future, the drivers would have the travel information offered by the route guidance system. Their formulation combined the user equilibrium with the available LP traffic control formulation at the time. In their problem statement, the goal was to determine the optimal metering rate so that the system measures of the network would be maximized, while the drivers would choose the path provided by the RG control. They discussed their mathematical formulation in detail and provided the solution finding algorithm.

In 1999 and then with some modifications in 2002, Apostolos Kotsialos et al. in [28, 29], considered the design of an integrated traffic control system for motorway networks with the use of ramp metering, motorway-to-motorway control, and route guidance. They offered a generic problem formulation in the format of a discrete time optimal control problem. They assumed that both RM control measures and RG are available. The METANET model was used for the description of traffic flow. A hypothetical test network was considered to evaluate the performance of the proposed control system. The control measure considered was the minimisation of the travel time spent (TTS). The results showed the high efficiency of the proposed control system.

In 2004, Karimi et al. [30] considered the integration of dynamic RG and RM based on MPC. They used the dynamic route guidance panels (DRIPs) as both a control tool and an information provider to the drivers, and ramp metering as a control tool to spread the congestion over the network. This resulted in a control strategy that reduced the total time spent by optimally re-routing traffic over the available alternative routes in the network, and also kept the difference between the travel times shown on the DRIPs and the travel times actually realized by the drivers as small as possible. The simulations done for the case study showed that rerouting of traffic and on-ramp metering using MPC has lead to a significant improvement in performance.

In 2015, Yu Han et al. [31] proposed an extended version of the CTM first proposed in [2] with the ability to reproduce the capacity drop at both the on-ramp bottleneck and the lane drop bottleneck. Based on this model, a linear quadratic model predictive control strategy for the integration of dynamic RG and RM was offered with the objective of minimizing the TTS of a traffic network. In this paper, a RG model based on the perceived travel time of each route by drivers was also offered. If the instantaneous travel time of each route is provided to travelers, the perceived travel time is assumed to be the same as the instantaneous travel time. If not, the perceived travel time is assumed to be the free flow travel time. The splitting rates at a bifurcation are determined by the well-known Logit model [26, 30]. A test case network containing both on-ramp bottlenecks and lane drop bottlenecks was used to investigate the effectiveness of the proposed framework and the results showed the improvement the proposed control strategy brought for the network performance.

In 2017, Cecilia Pasquale et al. [22] offered a multi-class control scheme for freeway traffic networks with the integration of RM and RG in order to reduce the TTS and the total emissions in a balanced way. Their two controllers were feedback predictive controllers and it was shown how this choice for their controllers can benefit the performance of the controllers. They applied the multi-class METANET model and the multi-class macroscopic VERSIT+ model for prediction of the traffic dynamics in the network. In addition, they designed a controller gain selector to compute the gains of the RM and RG controllers. The simulation results showed significant improvements of the freeway network performance, in terms of reduction of the TTS and the total emissions.

In 2018, Hirsh Majid et al. [32] designed integrated traffic control strategies for highway networks with the use of RG and RM. The highway network was simulated using the LWR model. A control algorithm was designed to solve the proposed problem, based on the inverse control technique and variable structure control (super twisting sliding mode). Three case studies were tested in the presence of an on-ramp at each alternate route and where there was a capacity constraint in the network. The objective was to avoid congestion on the main road and to balance the traffic flow on the alternate routes. The obtained results showed that the proposed algorithms could establish user equilibrium between two alternate routes even when the on-ramps have different traffic demands.

In 2019, Martin Gregurić et al. [33] proposed the approach of coordination between controlling on-ramp flows with ramp metering (RM) and dynamic route guidance information systems (DRGIS), which reroute vehicles from congested parts of the motorway. DRGIS is used to inform drivers about current or expected travel times and queue lengths so that they may reconsider their choice for a certain route. It can be seen that DRGIS can directly impact on traffic demand at the urban traffic system by informing the drivers about travel times on its crucial segments. Reduced traffic demand on congested urban motorway section or at congested on-ramp in coordination with the adequate ramp metering control strategies can prevent "spill-back effect" and increase overall throughput of the urban motorways.

2.5 Summary

To conclude, in this section, a review on the integration of RM and RG controllers was presented. The section started by describing the two most commonly used discrete first-order and second-order traffic flow models. Then, it continued with an overview of the popular RM and RG control strategies and it finished by discussing most of the important studies on the integration of RM and RG. Most of them considered TTS as the performance metric and applied an MPC optimization framework since it reduces the computation efforts required to solve the optimization problem specially if the traffic evolution model used was the METANET model since it makes the formulation non-linear and non-convex. Overall, all the studies reviewed here have shown improvements in the network performance in comparison with the case of having either of these controllers alone.

The discussion so far has centered around the macro-level control of intelligent highways. However, an integral component of the transportation system of a smart city is the micro-level control of autonomous vehicles. It is, therefore, imperative that we cover some micro-level details pertaining to the CAVs. The following sections address problems related to the micro-level vehicle coordination. Commonly found interactions include highway merging, off-ramp exit, vehicle overtaking and lane changing. Focus is placed on on-ramp merging and overtaking in presence of incoming traffic, as these two tasks combined encompass most of the complexities involved in inter-vehicle interactions.

The selection of the on-ramp merging task (see Section 3) as a key area to be explored is due to both the extent of variables involved in coordinating this process and the dynamic nature of the process itself. In fact this is one of the tasks that today's autonomous vehicles find difficult to carry out due to the need of reactive control and precise planning. The role of inter-vehicle coordination in efficient merging is also discussed in detail. While coordination of human-driven vehicles is mostly reactive, CAVs can be assigned goals proactively so as to optimize the merging process. Similarly, a detailed discussion is provided on the car overtake problem (see Section 4) to emphasize the need for explicit modelling of human behavior when designing algorithms for CAVs. This seemingly simple problem is specifically chosen to draw attention to the complexities that could arise due to the presence of human drivers on the road. A naive data-driven algorithm can fail catastrophically in scenarios where humans may behave unpredictably so an overview of algorithms that explicitly take human behavior into consideration is provided later on in this chapter.

3. Merging behavior at on-ramps and off-ramps for CAVs

In transportation networks, overall highway system efficiency can be severely reduced due to delays caused at on-ramps and off-ramps. If the merge process is incorrectly handled, merging lanes can have an overflow effect which causes the entire highway to become congested. This effect is caused by slower moving vehicles facing congestion in outer lanes near the merge junction deciding to switch into inner lanes in order to move faster. Therefore, even vehicles in the inner high speed lanes have to slow down. Overtime with continuing merge lane congestion, the entire freeway can become blocked.

In fact, in human-driven vehicles, this issue is further compounded due to the lack of cooperation and limited visibility for decision making. However, the introduction of CAVs has led to a lot more information becoming available for improving this overall merging process. In addition to the improved local sensing on-board modern CAVs such as 360° radar and vision based sensors, most of the increased information comes from improvements in Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communication methods. These communication protocols enable individual vehicles to broadcast their intent and status, receive command velocities to enable optimal flow and also collaborate with each other to self-organize in such a way that freeway traffic flow is optimized.

Effectively managing the highway merge problem has multiple benefits, both to the end users and the entire transportation system as well. Reduced time in traffic at merge junctions means that overall throughput of the highway is increased, individual waiting time is reduced and the wastage of fuel (energy) in idling vehicles stuck in traffic is also reduced. Therefore, it is evident that improvements to intelligent highways will have an impact on the economy as well as helping combat environmental issues caused by excessive fuel consumption.

3.1 Standard problem formulation

Most approaches to solving the highway merging problem use a similar structure to model the physical highway on-ramp. **Figure 2** shows an abstracted model of a highway on-ramp. A control zone is defined, where all vehicles in this zone communicate with a central controller and each other to decide individual optimum

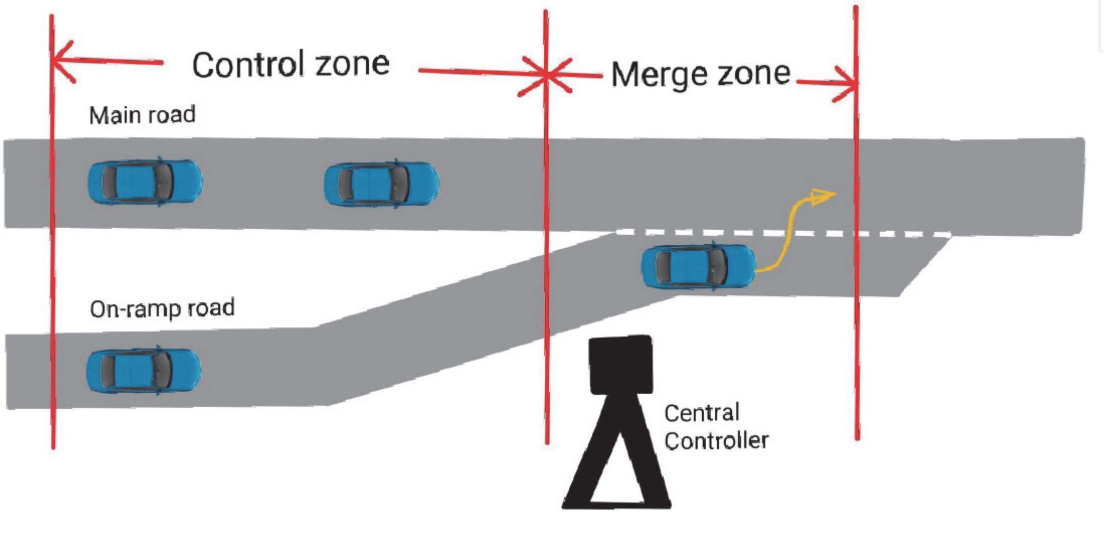

Figure 2.
On-ramp merging regions and infrastructure model.

velocities and paths to be followed. The control zone encompasses both the main road and the on-ramp. Vehicles are allowed to merge from the on-ramp onto the main road in the merge zone, which is located at the end of the control zone.

The vehicles involved are modelled based on simplistic second order dynamics given by,

$$\begin{aligned} \dot{p}_i &= v_i(t) \\ \dot{v}_i &= u_i(t) \end{aligned} \tag{32}$$

where $p_i(t)$, $v_i(t)$, $u_i(t)$ denotes the position, velocity and acceleration/deceleration (control input) respectively for each vehicle. The vehicle state is then defined as,

$$x_i(t) = \begin{bmatrix} p_i(t) \\ v_i(t) \end{bmatrix} \tag{33}$$

Furthermore, assuming that lateral control keeping the vehicle in lane is managed elsewhere, the vehicle can be modelled as a point mass moving along the center of the lane with the following state equation, where time 0 is the point at which the vehicle enters the control zone.

$$\dot{x}_i = f(t, x_i, u_i), \qquad x_i\left(t_i^0\right) = x_i^0 \tag{34}$$

Additionally, all the methods discussed have the shared assumption that, *the vehicle speed inside the merging zone is constant*.

To compare these algorithms, the two main performance indicators are *throughput* (maximum number of vehicles that can merge onto the highway in an hour) and *delay* (average delay experienced by vehicles compared to the ideal travel time). In addition to these parameters, some research in this area also takes into consideration the savings in fuel consumption due to improvements in the highway merging process.

3.2 Various methods

So, let's now explore some of the methods used in handling the highway merging problem in greater detail. Here, multiple approaches and methodologies to address this problem are discussed. Most of the work done in finding an optimal solution to automated freeway merging is based on posing the problem in the form of an optimization problem [34] with centralized control [35], virtual slot-based dynamics [36] problem or as broadcast communication [37] problem.

3.2.1 Optimal control method

This problem is posed as an unconstrained optimization problem [34] and then further extended to also consider the impact of fuel consumption [35]. The output was the ability to derive online an optimal closed-form solution for vehicle coordination at a merge intersection. The importance of safety constraints is also stressed here. Constraints in positioning, maximum velocity and maximum accelerations are imposed. Additionally, the algorithm only allows one vehicle to enter into the merging zone at any time.

The algorithm calculates the time at which each vehicle would enter the merge zone and requires that this time does not conflict with any of the other vehicles.

This ensures that a lateral collision can never happen. Additionally, the algorithm also directs that vehicles maintain at least a specified gap between each other which ensures that rear-end collisions do not occur. Hamiltonian analysis is then used to convert the optimization problem into a system of four equations that can be solved in real time to output the optimal control for each vehicle.

Simulation of this system was then carried out to show that the algorithm performs as desired. It was found that compared to a baseline situation where on-ramp vehicles always give way to vehicles on the freeway this algorithm performs significantly better. An improvement of 52% in fuel consumption when compared to the baseline situation was also reported.

Disadvantages: Only one lane of the freeway is in use and the benefits obtained from allowing/forcing vehicles to switch lanes in the freeway are ignored (i.e. full capacity of the freeway is not used). Additionally, vehicles are given merging rights based on a simplistic FIFO (First In First Out) queue which can cause additional delays and is definitely sub-optimal.

3.2.2 Slot based method

This method [36] primarily relies on creating virtual slots for each vehicle that moves along the freeway at a constant velocity. Then all changes to this behaviour such as switching lanes, on-ramp merging and exiting the highway on an off ramp are modelled as a switch from one virtual slot to another. A virtual slot S is defined with five properties as is denoted by $S = \{z, p, t, b, o\}$, where z is the size of the slot, p is the position, t is the time, b is the behaviour of the slot and o is the density status of the slot.

Slots are created by a central slot controller and vehicles can request to change from one slot to another. This change will then be approved by the slot controller as long as the slot is not already occupied or there is no other vehicle requesting to switch to that same slot. This approach was also shown to work in the absence of infrastructure at the merge junction since vehicles can use V2V communication to find an unoccupied virtual slot and perform the merging task.

An additional benefit of this slot based system is the ease by which the method can be extended to allow the entire bandwidth of the freeway to be used in order to further improve efficiency. For example, if the slot controller realizes that there are a lot of empty slots in the central lanes of the freeway, the controller can request that vehicles in the outer lane prior to the merge point to move into the empty inner lane slots. This creates more empty slots in the outer lane and provides more opportunities for vehicles on the on-ramp to merge successfully. This type of cooperative behaviour has been proven to drastically improve the throughput of vehicles through these freeway merge zones.

Using simulations, it has been shown that throughput at merge intersections can be increased and delay can be decreased drastically (throughput: 230% increase and delay: 452% decrease) vs. human driven vehicles under heavy traffic conditions.

Disadvantages: Many limitations brought about by having slot based systems include, difficulty in robustly handling emergency/breakdown situations, lack of flexibility in catering to different needs such as different vehicles requesting different speeds, inefficiency in heavy traffic density situations where not enough free slots are available to facilitate lane changes etc. Moreover, the slot based method places a lot of restrictions on the way vehicles can move about and position themselves on the freeway. Communication between vehicles and infrastructure also needs to be extremely good for this system to work and this type of perfect communication is rarely available in practice.

3.2.3 Broadcast communication based method

Some of the major issues in coordinating automated vehicles at merge intersections are the problems caused by imperfect communication. In these type of applications, delays of even a few seconds can have devastating results. Work on using Pseudo-perturbation based broadcast communication (PBC) [37] instead of unicast communication focuses on reducing the overhead on the V2I communication systems and leveraging the capabilities of V2V communication to handle any short term changes. In this method, a global controller (infrastructure) broadcasts an identical message to all vehicles in its vicinity. Each of these vehicles uses this information along with V2V data from other vehicles to select a suitable control strategy to safely perform the coordinated merging task. Vehicles send updates back to the central controller and the controller uses this feedback to decide its next broadcast message.

The key focus of this research was to extend available PBC capabilities to handle the multi-state vehicle dynamics and the multi-objective optimization required to solve a freeway merging coordination problem. The capability of this system to handle both CAVs as well as a mix of CAVs and human driven vehicles was also showcased. Here, the CAVs are controlled by the coordination system and are shown to be able to function even in the presence of human driven vehicles. The smoothness of the actual merging process was also evaluated through simulation. The main output of this research was to show that complex coordination problems such as freeway merging can be successfully solved with the use of minimal communication bandwidth.

Disadvantages: Lack of focus on the actual merging algorithm. Based on the broadcast signals, the decisions individual vehicles make may be sub-optimal and cause an unnecessary delay in the system. Also, very little work has been done on seeing whether this system actually has a effect on throughput.

3.2.4 Temporal logic based method

This method involves formalizing traffic rules using temporal logic [38] in order to ensure safety and robustness of automated highway vehicle control. This method helps with formulating existing traffic rules in a mathematical way in order to be easily applied in CAVs. While this method does not solely focus on the merging problem, the merge window is addressed in the metric temporal logic (MTL) formulas included. When the merge operation is formulated as a MTL formula, it leads the way to specifying safety guarantees in autonomous vehicles. Furthermore, the legality of trajectories generated by a motion planner can be easily checked using these MTL formulae. It also allows the use of standard verification and validation methods in order to ensure that there are no loopholes or issues in the generated logic.

3.3 Summary

While there are many methods to improve intelligent merging behavior, the core fundamentals of these algorithms are quite similar. They look into minimizing gaps or under-utilized space on the road while minimizing the control inputs (acceleration and braking) needed to achieve this. All algorithms also prioritize safety and fairness in the merging process. Additionally, there are extensions that focus on maximum capacity utilization by moving vehicles already on the highway to less congested lanes, which further improves the efficiency of the merging process. Each of the algorithms discussed in the section has its own advantages and disadvantages. Therefore, it falls on the highway regulatory agencies to decide which of these are most suitable to the conditions of each individual highway system.

4. The overtaking behavior with the combination of HDVs and CAVs

The final section of this chapter focuses on providing an in-depth analysis of a complex scenario in which the autonomous vehicle (AV) has to perform maneuvers in the presence of HDVs. Upon a cursory glance at the current state of AV research [39], it becomes obvious that not a lot of emphasis is being placed on explicitly modelling the varying behavior patterns of HDVs on the road. One such instance that highlights the need to model the varying HDV driving patterns is the car overtaking problem in a bidirectional traffic flow setting. In this problem, a scenario with three vehicles is considered; two HDVs and one AV (ego vehicle), as shown in **Figure 3**. The HDVs are travelling in opposite directions in adjacent lanes and the ego vehicle is following one of the HDVs. The objective of the AV is to safely overtake the vehicle travelling ahead while maintaining safety distances to the HDV in the adjacent lane, the HDV travelling ahead and the boundaries of the road.

This is a particularly hard problem to solve because it involves a scenario composed of both human-driven and autonomous vehicles. The first major complication is the lack of global information because the V2X communication protocols cannot be leveraged in this scenario due to the presence of HDV. Then, there is the problem of uncertainty that arises due to the varying driving patterns so the algorithm needs to be robust enough to handle the different driving patterns of human drivers. Moreover, the traditional Supervised Learning based approaches cannot be applied directly to this problem due to a lack of labeled training data. Finally, the model-free Reinforcement Learning (RL) based techniques [40] cannot be employed due to a lack of safety guarantees while the model-based RL or Control techniques [41] cannot be employed since they require an accurate representation of system model and cannot capture uncertainties that arise due to varying driving patterns well.

In this chapter, a simplified stochastic control based formulation taken from [42] is laid out to provide a mathematical description of the problem. The formulation is followed by a brief discussion of a couple of algorithms to give the readers a glimpse of the possible avenues that could be taken to reach a solution. The references to detailed resources are also provided for the interested readers to explore further.

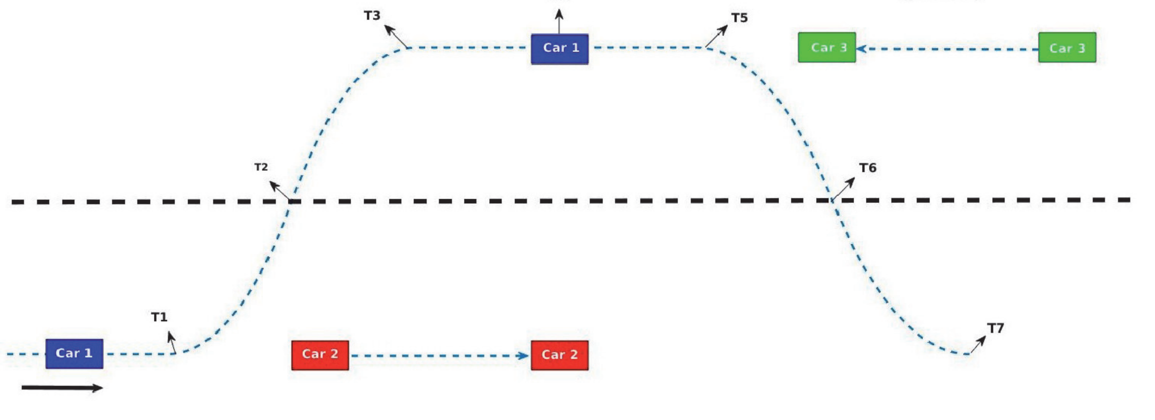

Figure 3.
Overview of the car overtaking problem.

4.1 Stochastic control formulation

In this approach, the ego vehicle (Car 1, from **Figure 3**) has to first decide if it is feasible to overtake the HDV ahead (Car 2) once it gets "close enough". In this simplified formulation, it is assumed that the Car 1 is able to measure its own relative velocity with respect to the HDVs with some additive noise. If the decision to overtake is made, then the AV has to generate the exact trajectory it will take to perform the overtake maneuver. An overview of these steps is outlined henceforth.

In the formulation below, it is assumed that the width of each of the lanes is defined to be a constant value d and the minimum safety distance between cars is L. A collision is defined in terms of AV violating the minimum safety distance threshold with respect to the centroids of the cars. The positive velocities are defined towards the right in **Figure 3** and positive θ is defined counterclockwise relative to the velocity vector of the car. Finally, the sets of admissible linear and angular speeds for the cars are considered to be finite.

4.1.1 Modelling

The states of Car i are x_i, y_i and θ_i which respectively correspond to the longitudinal coordinate, lateral coordinate and orientation of Car i. The inputs to Car i are v_i and w_i, the linear and angular velocities respectively of Car i respectively. An index k is used to denote the time step.

Considering the initial states of the vehicles, it is assumed that initially at $k = 0$, Car 1's longitudinal coordinate is a random variable distributed normally with $\mu = 0$ and $\sigma = \Sigma_1$ while the lateral coordinate is fixed at the center of the bottom lane i.e. $y_1(0) = d/2$ and facing forward i.e. $\theta = 0$. As for Car 2, it is assumed that initially at $k = 0$, x_2 is distributed normally with $\mu_2 = \tilde{x}_2$ and $\sigma = \Sigma_2$ where Σ_2 is s.t. $x_2(0) > x_1(0)$ while the lateral coordinate is at $y_2(0) = d/2$ and facing forward i.e. $\theta = 0$, identical to Car 1. Finally, for Car 3, it is assumed that initially at $k = 0$, x_3 is distributed normally with $\mu_3 = \tilde{x}_3$ and $\sigma = \Sigma_3$ where Σ_3 is s.t. $x_3(0) > x_2(0)$ while the lateral coordinate is fixed at the center of the top lane i.e. $y_3(0) = 3d/2$ and facing reverse (since $v_3 < 0$) i.e. $\theta = 0$.

As for the dynamics of the vehicles, it is assumed that cars 2 and 3 keep travelling along the same lane i.e. ω_2 and ω_3, the angular velocities of cars 2 and 3 respectively, are identically zero for all time steps $k \geq 0$.

Based on the assumptions above, the dynamics for the cars are defined by the equations below:

$$x_1(k+1) = x_1(k) + v_1(k)\cos(\theta_1(k)) \tag{35}$$

$$y_1(k+1) = y_1(k) + v_1(k)\sin(\theta_1(k)) \tag{36}$$

$$\theta_1(k+1) = \theta_1(k) + \omega(k) \tag{37}$$

$$x_2(k+1) = x_2(k) + v_2(k) \tag{38}$$

$$y_2(k) = d/2 \tag{39}$$

$$\theta_2(k) = 0 \tag{40}$$

$$x_3(k+1) = x_3(k) + v_3(k) \tag{41}$$

$$y_3(k) = 3d/2 \tag{42}$$

$$\theta_3(k) = 0 \tag{43}$$

To keep the problem as general as possible, it is assumed that at time step k, the AV has all the history of its past states, linear velocity and its relative position and velocity w.r.t the other two cars with some additive white Gaussian noise (AWGN). Therefore,

$$\mathcal{I}(k) = \{x_1(n), y_1(n), \theta_1(n), v_1(n), z_1(n), z_2(n), z_3(n), z_4(n)\}_{n=0}^{n=k} \tag{44}$$

where

$$z_1(k) = x_2(k) - x_1(k) + n_1(k) \tag{45}$$

$$z_1(k) = x_3(k) - x_1(k) + n_2(k) \tag{46}$$

$$z_1(k) = v_2(k) - v_1(k) + n_3(k) \tag{47}$$

$$z_1(k) = v_3(k) - v_1(k) + n_4(k) \tag{48}$$

Here, n_i(k) are white Gaussian Processes with mean 0 and variances $\sigma_i^2(k)$. There is an assumption placed on the independence of $\{n_i(k)\}_{k\geq 1}$ from the distribution of initial longitudinal coordinates of the three cars i.e. $x_1(0)$, $x_2(0)$ and $x_3(0)$.

4.1.2 Control problem

Upon consideration of the control problem that the AV has to solve, it becomes apparent rather quickly that the AV simply doesn't have to decide on linear and angular velocities. If that were the case, then the Car 1 could simply wait till Car 3 passes and then overtake Car 2 by having a velocity greater than Car 2. That is not an optimal solution for all possible scenarios. Therefore,the first thing that the AV needs to do is to get better estimates of position and velocity of Cars 2 and 3 rather than using the raw noisy data. With the better estimates of positions and velocities of the HDVs, the AV can perform a feasibility analysis to see if it is feasible to overtake the Car 2 or not. If the Car 1 deems the overtake maneuver to be infeasible, it can resort to the waiting strategy. If, however, Car 1 decides to overtake, then it has to decide on the time to start overtaking. Once the time to start overtaking is finalized, then the AV has to plan a trajectory that it will take to perform the maneuver such that it will not violate any safety margins. Finally, it needs to generate control commands i.e. linear and angular velocities to execute the overtaking maneuver. One typical problem that could arise is that the vehicle may have to return to its original lane, after starting the overtaking maneuver, but this scenario is beyond the discussion of this chapter.

4.1.3 Sample trajectory

A sample trajectory with a constant linear velocity for the AV is displayed in **Figure 4**. Between $k = 0$ and $k = T1$, Car 1 is approaching Car 2. At $k = T1$, Car 1 decides to run the feasibility analysis and decides to overtake Car 2. Between $k = T1$ and $k = T2$, Car 1 has a positive constant angular velocity resulting in motion towards the adjacent lane. At $k = T2$, Car 1 has reached the divider between the lanes so it switches to a negative angular acceleration having the same magnitude as

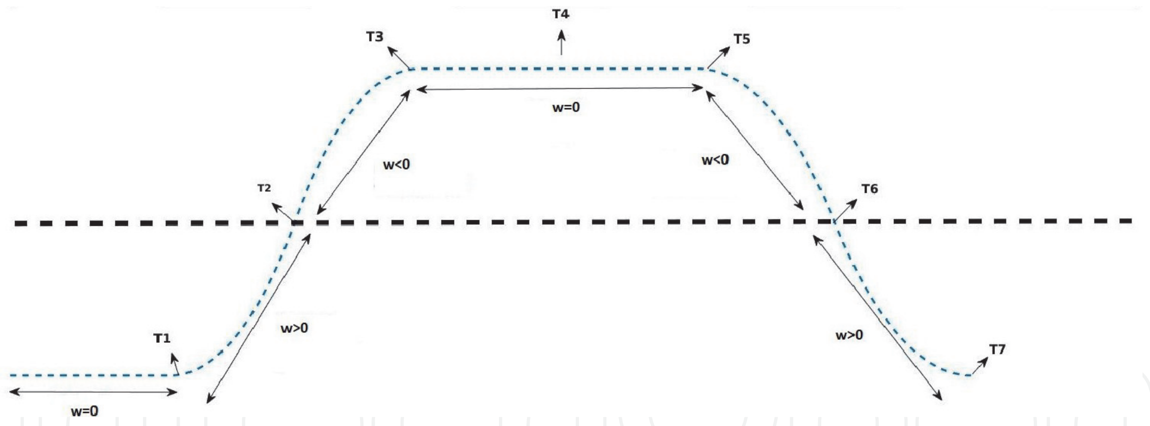

Figure 4.
Sample trajectory for car overtaking problem.

before, until the car reaches the center of adjacent lane at $k = T3$. Between $k = T3$ and $k = T5$, the angular acceleration remains at 0 for a straight line motion in opposite lane for overtaking. A negative angular acceleration having the same magnitude as before is applied between $k = T5$ and $k = T6$ followed by positive angular acceleration between $k = T6$ and $k = T7$. Car 1 continues its motion in a straight line after $k = T7$ with zero angular acceleration.

4.2 Solution methods

There were quite a few assumptions made to obtain the simplified model discussed above and some of those assumptions could be removed in order to obtain a rather complicated yet general framework. For instance, it might not be possible for the AV to store all the history of past states and actions so an attempt at modelling with limited information could be made. With this disclaimer, a brief overview of some of the possible solution methods for this problem is presented below.

4.2.1 Minimizing probability of collision

In this approach, taken from [42], it is assumed that the AV will travel at a constant speed throughout the overtaking maneuver. The estimates of positions and velocities of HDVs are obtained using Kalman filtering and a constant N is introduced to characterize the behavior of the AV with higher values corresponding to higher level of aggressiveness. Two feasible sets for the linear and angular velocities respectively of Car 1 are obtained by ensuring that the estimated position of AV after performing the maneuver will stay outside the estimated minimum safety region around the HDVs. Using the feasibility sets, probability of collision of Car 1 with the HDVs is obtained and the decision to overtake is based on those probabilities. If the decision to overtake is made, the linear and the angular velocities are chosen to minimize the probability of collision and the maneuver is performed as detailed in Section 4.1.3.

4.2.2 Reachability analysis-based with martingale-based HDV modelling

In this approach, taken from [43], the focus is on obtaining safety guarantees while overtaking. In this approach, the restriction on the constant speed of HDVs is also lifted, which was alluded to previously in Section 4.2. There are two different reachability analysis-based algorithms presented: one is a robust time-optimal algorithm such that it provides strict guarantees in regards to collision avoidance while the other is a stochastic algorithm that that yields a small collision probability with the advantage of shorter overtaking time. Moreover, the expected behavior of the human driver is modeled using a stochastic model based on martingales. It is shown that if the human driver is non-aggressive, the stochastic algorithm will yield a shorter overtaking time and if the driver is aggressive, the behaviors for stochastic and the robust algorithms will be identical.

4.3 Summary

The key takeaway from this section is that there is a need to place focus on explicitly considering the role of human drivers on the road while developing algorithms for autonomous vehicles. It was shown by the study of the simple car overtaking example that there are scenarios where the need to model human drivers on the road increases manifold. The algorithm presented in Section 4.2.2 explicitly models the behavior of human-drivers with martingales and provides overtaking algorithms with safety guarantees. The solution complexity of this approach is rather high due to the reachability analysis-based solutions so further research could be directed at improving the solution complexity or coming up with approaches that yield lower complexity while maintaining the safety guarantees. Furthermore, there is a prospect of exciting research in the direction of modeling human behavior with other approaches which may lead to various other interesting algorithms. The aim of this section was to provide motivation to the reader to explore research avenues that incorporate explicit modelling of human driving patterns yielding algorithms that will expedite the introduction of Autonomous Vehicles onto our roads.

5. Conclusions

The advent of improved communication, sensing and control in modern day vehicles and infrastructure creates a lot of opportunities to improve the efficiency and safety in many highway processes. Key areas of interest involve traffic routing and management, optimal highway merging and intelligent overtaking behaviours. This chapter examined some of the methods used in these areas and discussed the various improvements and shortcomings of each of them. The implementation of the algorithms discussed in this chapter, would lead to modern transportation systems becoming more effective, productive and safe. While there are many other methods worthy of merit not discussed in this chapter, the areas covered should give the reader a broad understanding of the extent of possibilities in this field and also spark further thinking which may lead to the generation of innovative new solutions.

Author details

Fatemeh Alimardani, Nilesh Suriyarachchi, Faizan M. Tariq
and John S. Baras*
Department of Electrical and Computer Engineering, Institute for Systems
Research, University of Maryland, College Park, Maryland, USA

*Address all correspondence to: baras@umd.edu

References

[1] Lighthill MJ, Whitham GB. On kinematic waves I. Flood movement in long rivers. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences. 1955 May 10;**229**(1178):281-316

[2] Daganzo C. The cell transmission model Part I: a simple dynamic representation of highway traffic. PATH Report, 93-0409. 1993;3.

[3] Papageorgiou M, Blosseville JM, Hadj-Salem H. Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris. Transportation Research Part B: Methodological. 1989 Feb 1;**23**(1):29-47

[4] Lu XY, Qiu TZ, Horowitz R, Chow A, Shladover S. METANET model improvement for traffic control. In 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC) 2011 Oct 5 (pp. 2148-2153). IEEE.

[5] Ferrara A, Sacone S, Siri S. Freeway traffic modelling and control. Berlin: Springer; 2018 Apr 12

[6] Gomes G, Horowitz R. Optimal freeway ramp metering using the asymmetric cell transmission model. Transportation Research Part C: Emerging Technologies. 2006 Aug 1; **14**(4):244-262

[7] Ferrara A, Sacone S, Siri S. Event-triggered model predictive schemes for freeway traffic control. Transportation Research Part C: Emerging Technologies. 2015 Sep 1;**58**:554-567

[8] Srivastava A, Geroliminis N. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. Transportation Research Part C: Emerging Technologies. 2013 May 1;**30**: 161-177

[9] Roncoli C, Papageorgiou M, Papamichail I. Traffic flow optimisation in presence of vehicle automation and communication systems–Part I: A first-order multi-lane model for motorway traffic. Transportation Research Part C: Emerging Technologies. 2015 Aug 1;**57**: 241-259

[10] Daganzo CF. The cell transmission model, part II: network traffic. Transportation Research Part B: Methodological. 1995 Apr 1;**29**(2):79-93

[11] Muralidharan A, Dervisoglu G, Horowitz R. Freeway traffic flow simulation using the link node cell transmission model. In2009 American Control Conference 2009 Jun 10 (pp. 2916-2921). IEEE.

[12] Canudas-de-Wit C, Ferrara A. A Variable-Length cell road traffic model: Application to ring road speed limit optimization. In2016 IEEE 55th Conference on Decision and Control (CDC) 2016 Dec 12 (pp. 6745-6752). IEEE.

[13] Messner A, Papageorgiou M. METANET: A macroscopic simulation program for motorway networks. Traffic engineering & control. 1990;**31** (8-9):466-470

[14] Pasquale C, Sacone S, Siri S. Two-class emission traffic control for freeway systems. IFAC Proceedings Volumes. 2014 Jan 1;**47**(3):936-941

[15] Pasquale C, Sacone S, Siri S, De Schutter B. A multi-class model-based control scheme for reducing congestion and emissions in freeway networks by combining ramp metering and route guidance. Transportation Research Part C: Emerging Technologies. 2017 Jul 1;**80**: 384-408

[16] Zhang M, Kim T, Nie X, Jin W, Chu L, Recker W. Evaluation of on-ramp control algorithms.

[17] Papageorgiou M, Kotsialos A. Freeway ramp metering: An overview. IEEE transactions on intelligent transportation systems. 2002 Dec;**3**(4): 271-281

[18] Sanandaji A, Ghanbartehrani S, Mokhtari Z, Tajik K. A Novel Ramp Metering Approach Based on Machine Learning and Historical Data. arXiv preprint arXiv:2005.13992. 2020 May 26.

[19] Papageorgiou M, Hadj-Salem H, Blosseville JM. ALINEA: A local feedback control law for on-ramp metering. Transportation Research Record. 1991 Jan;**1320**(1):58-67

[20] Wang Y, Kosmatopoulos EB, Papageorgiou M, Papamichail I. Local ramp metering in the presence of a distant downstream bottleneck: Theoretical analysis and simulation study. IEEE Transactions on Intelligent Transportation Systems. 2014 Mar 31; **15**(5):2024-2039

[21] Wardrop JG. Road paper. some theoretical aspects of road traffic research. Proceedings of the institution of civil engineers. 1952 Jun;**1**(3):325-362

[22] Pasquale C, Sacone S, Siri S, De Schutter B. A multi-class model-based control scheme for reducing congestion and emissions in freeway networks by combining ramp metering and route guidance. Transportation Research Part C: Emerging Technologies. 2017 Jul 1;**80**: 384-408

[23] Pavlis Y, Papageorgiou M. Simple decentralized feedback strategies for route guidance in traffic networks. Transportation science. 1999 Aug;**33**(3): 264-278

[24] Mahmassani HS, Peeta S. Network performance under system optimal and user equilibrium dynamic assignments: implications for ATIS. Transportation Research Board; 1993 Jan 10.

[25] Wang Y, Messmer A, Papageorgiou M. Freeway network simulation and dynamic traffic assignment with METANET tools. Transportation Research Record. 2001; **1776**(1):178-188

[26] Theil H. A multinomial extension of the linear logit model. International Economic Review. 1969;**10**(3):251-259

[27] Iida Y, Hasegawa T, Asakura Y, Shao CF. A formulation of on-ramp traffic control system with route guidance for urban expressway. IFAC Proceedings Volumes. 1990 Sep 1;**23**(2): 229-236

[28] Kotsialos A, Papageorgiou M, Messner A. Integrated optimal control of motorway traffic networks. In Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251) 1999 Jun 2 (Vol. 3, pp. 2183-2187). IEEE.

[29] Kotsialos A, Papageorgiou M, Mangeas M, Haj-Salem H. Coordinated and integrated control of motorway networks via non-linear optimal control. Transportation Research Part C: Emerging Technologies. 2002 Feb 1; **10**(1):65-84

[30] Karimi A, Hegyi A, De Schutter B, Hellendoorn H, Middelham F. Integration of dynamic route guidance and freeway ramp metering using model predictive control. In Proceedings of the 2004 American Control Conference (ACC 2004) 2004 Jun 30 (pp. 5533-5538).

[31] Han Y, Yuan Y, Hegyi A, Hoogendoorn SP. Linear quadratic MPC for integrated route guidance and ramp metering. In2015 IEEE 18th International Conference on Intelligent Transportation Systems 2015 Sep 15 (pp. 1150-1155). IEEE.

[32] Majid H, Lu C, Karim H. An integrated approach for dynamic traffic

routing and ramp metering using sliding mode control. Journal of Traffic and Transportation Engineering (English Edition). 2018 Apr 1;**5**(2):116-128

[33] Gregurić M, Mandžuka S, Vidović K. Advanced Applications for Urban Motorway Traffic Control. InInternational Conference "New Technologies, Development and Applications" 2019 Jun 27 (pp. 395-400). Springer, Cham.

[34] Rios-Torres J, Malikopoulos A, Pisu P. Online optimal control of connected vehicles for efficient traffic flow at merging roads. In 2015 IEEE 18th international conference on intelligent transportation systems 2015 Sep 15 (pp. 2432-2437). IEEE.

[35] Rios-Torres J, Malikopoulos AA. Automated and cooperative vehicle merging at highway on-ramps. IEEE Transactions on Intelligent Transportation Systems. 2016 Aug 5; **18**(4):780-789

[36] Marinescu D, ÄŒurn J, Bouroche M, Cahill V. On-ramp traffic merging using cooperative intelligent vehicles: A slot-based approach. In2012 15th International IEEE Conference on Intelligent Transportation Systems 2012 Sep 16 (pp. 900-906). IEEE.

[37] Ito Y, Kamal MA, Yoshimura T, Azuma SI. Multi-vehicle coordination on merging roads based on pseudo-perturbation-based broadcast control. In2018 Annual American Control Conference (ACC) 2018 Jun 27 (pp. 4008-4013). IEEE.

[38] Maierhofer S, Rettinger AK, Mayer EC, Althoff M. Formalization of Interstate Traffic Rules in Temporal Logic. Proc. of the IEEE Intelligent Vehicles Symposium. 2020

[39] Badue C, Guidolini R, Carneiro RV, Azevedo P, Cardoso VB, Forechi A, Jesus L, Berriel R, Paixao T, Mutz F, Veronese L. Self-driving cars: A survey. arXiv preprint arXiv:1901.04407. 2019 Jan 14.

[40] Li X, Xu X, Zuo L. Reinforcement learning based overtaking decision-making for highway autonomous driving. In2015 Sixth International Conference on Intelligent Control and Information Processing (ICICIP) 2015 Nov 26 (pp. 336-342). IEEE.

[41] Petrov P, Nashashibi F. Modeling and nonlinear adaptive control for autonomous vehicle overtaking. IEEE Transactions on Intelligent Transportation Systems. 2014 Mar 3; **15**(4):1643-1656

[42] Raghavan A, Wei J, Baras JS, Johansson KH. Stochastic control formulation of the car overtake problem. IFAC-PapersOnLine. 2018 Jan 1;**51**(9):124-129

[43] Gao Y, Jiang FJ, Johansson KH, Xie L. Stochastic Modeling and Optimal Control for Automated Overtaking. In 2019 IEEE 58th Conference on Decision and Control (CDC) 2019 Dec 11 (pp. 1273-1278). IEEE.