Transportation Research Part C 44 (2014) 242–252

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Real-time merging traffic control for throughput maximization at motorway work zones

A. Tympakianaki [a,*], A. Spiliopoulou [a], A. Kouvelas [a], I. Papamichail [a], M. Papageorgiou [a], Y. Wang [b]

[a] Dynamic Systems and Simulation Laboratory, Technical University of Crete, 73100 Chania, Greece
[b] Institute of Transportation Engineering, College of Civil Engineering and Architecture, ZheJiang University, Hangzhou 310058, PR China

ARTICLE INFO

Article history:
Received 9 February 2014
Received in revised form 15 April 2014
Accepted 15 April 2014

Keywords:
Work zone management
Feedback control
Merging traffic control
Adaptive fine-tuning (AFT)
Regulator fine-tuning

ABSTRACT

Work zones on motorways necessitate the drop of one or more lanes which may lead to significant reduction of traffic flow capacity and efficiency, traffic flow disruptions, congestion creation, and increased accident risk. Real-time traffic control by use of green–red traffic signals at the motorway mainstream is proposed in order to achieve safer merging of vehicles entering the work zone and, at the same time, maximize throughput and reduce travel delays. A significant issue that had been neglected in previous research is the investigation of the impact of distance between the merge area and the traffic lights so as to achieve, in combination with the employed real-time traffic control strategy, the most efficient merging of vehicles. The control strategy applied for real-time signal operation is based on an ALINEA-like proportional–integral (PI-type) feedback regulator. In order to achieve maximum performance of the control strategy, some calibration of the regulator's parameters may be necessary. The calibration is first conducted manually, via a typical trial-and-error procedure. In an additional investigation, the recently proposed learning/adaptive fine-tuning (AFT) algorithm is employed in order to automatically fine-tune the regulator parameters. Experiments conducted with a microscopic simulator for a hypothetical work zone infrastructure, demonstrate the potential high benefits of the control scheme.

1. Introduction

Work zone management aims at safe working conditions for work-zone workers, as well as safe and efficient passage of vehicles. An extensive report by the FHWA (2005) provides very useful insights on motivation, traffic management possibilities and impact. This paper addresses the sub-class of major motorway work zones, where one or more lanes need to be closed over a period of several days or months; as a consequence, the traffic flow approaching the work zone needs to merge from a higher number of lanes into a lower number of lanes within a limited space. When the arriving flow reaches or exceeds the reduced downstream capacity, congestion is created, leading to an additional, congestion-induced capacity drop; although the exact reasons for the occurrence of capacity drop have not been fully explored until now, a major

* Corresponding author. Present address: Traffic and Logistics Division, KTH Royal Institute of Technology, Teknikringen 72, 100 44 Stockholm, Sweden. Tel.: +46 8 7908891.

E-mail addresses: athina.tymp@gmail.com (A. Tympakianaki), natasa@dssl.tuc.gr (A. Spiliopoulou), tasos.kouvelas@gmail.com (A. Kouvelas), ipapa@dssl.tuc.gr (I. Papamichail), markos@dssl.tuc.gr (M. Papageorgiou), wangyibing@zju.edu.cn (Y. Wang).

http://dx.doi.org/10.1016/j.trc.2014.04.006
0968-090X/

influencing factor is deemed to be due to the need for vehicles to accelerate from low speeds within the congestion to higher speeds downstream of the congestion head.

For example, referring to a tunnel bottleneck, it was found already 50 years ago that "*congestion inside the tunnel reduces its throughput. ... due to the fact that most cars do not accelerate very efficiently once they have to stop, or even just slow down*" (Gazis and Foote, 1969), see also the discussion in Papageorgiou et al. (2008)). Also, empirical field experiments carried out in Japanese motorways at sag bottlenecks indicate that the capacity drop is mitigated if drivers are alerted to accelerate promptly at the head of the congestion (Murashige, 2011).

In the past, several procedures and strategies have been proposed or used to improve traffic conditions at work zones, including speed limitations, as well as signing, markings and particular geometric design (see e.g. Lin et al., 2004; FHWA, 2005; Wei and Pavithran, 2006). More recently, real-time merging traffic control was proposed (Lentzakis et al., 2008), aiming at throughput maximization and minimization of delays in work zones in a similar way as the mainstream traffic flow control concept by Carlson et al. (2010), albeit by use of traffic lights instead of variable speed limits. As a matter of fact, mainstream traffic flow control was already applied in the late 1950s and 1960s to increase the throughput of the tunnels under the Hudson River, which connect New York City with New Jersey (Gazis and Foote, 1969). More recent work on mainstream traffic flow control is reviewed in Carlson et al. (2011).

This paper continues on the work of Lentzakis et al. (2008) and Papageorgiou et al. (2008) by improving and extending the related investigations, tools and insights in a number of ways which increase the chances of a successful field deployment of the method. To start with, an important issue that had been neglected in previous works is the impact of the positioning of the traffic lights when applying real-time traffic control. The distance between the traffic lights and the merge area is crucial as it affects the vehicles' behavior and particularly the acquired speed when approaching the merge area. It is shown in this paper that the appropriate location of the traffic lights may improve the results of merging traffic control, as the capacity drop can be mitigated or even eliminated in case of proper merging vehicle speeds, and this contributes to a more efficient and safe passage through the merge area. This aspect is farther highlighted by the inclusion of trucks in the simulation-based demonstration. In view of the potentially longer distance of the traffic lights from the merge area, the usage of a PI-type feedback regulator (in replacement of the previously used I-type regulator) is employed in this paper.

Another novel issue addressed, is the calibration of the employed regulator parameters for the applied control strategy by use of a recently proposed (Kosmatopoulos, 2009; Kouvelas, 2011; Kouvelas et al., 2011) automatic fine-tuning procedure (called AFT). Novel aspects in this respect include the usage of both Support Vector Machine (SVM) and polynomial, approximators within AFT, along with the comparison of the corresponding resulting tuning behaviors; as well as the specific fine-tuning results for simple PI-regulators and their set-points, along with the investigation of application conditions that ensure improved performance of the utilized control strategy. This fine-tuning investigation is, at the same time, an additional demonstration of the automatic fine-tuning capabilities of AFT, the relevance of which reaches beyond the present work zone traffic control application.

2. Work zone traffic control

2.1. Merge area

A typical motorway work zone area is sketched in Fig. 1a. The vehicles arriving on M lanes, must change lanes appropriately within the (typically trapezoidal) merge area (or earlier) so as to fit into the μ lanes of the exit (where M is higher than μ). The merging procedure may be quite complex in terms of the required vehicle maneuvers, especially when the traffic density in the merge area is high.

The capacity of work zone areas is usually lower than the mainstream motorway capacity due to the drop of one or more lanes at the work zone entrance. Fig. 1b displays a typical flow-density diagram for the merge area, where the flow q_{out} is the

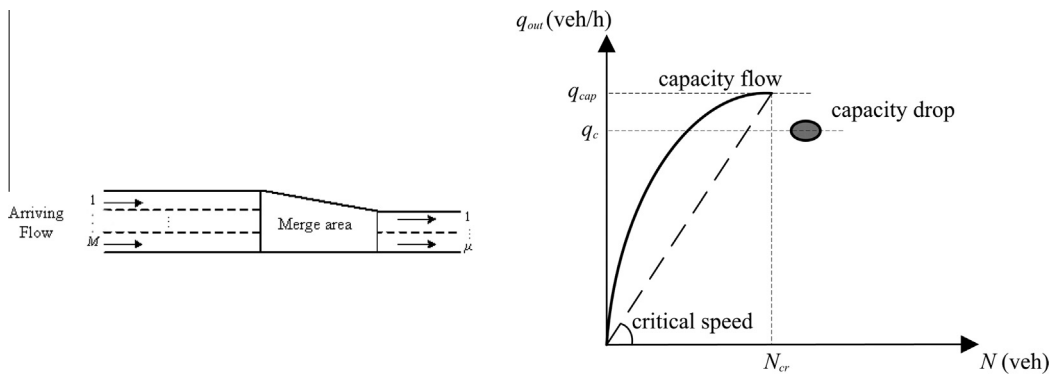

Fig. 1. (a) Typical motorway work zone area and (b) fundamental diagram of a merge area.

merge area exit flow and N is the number of vehicles included in the merge area. When the number of vehicles N included in the merge area is small, merging conflicts are scarce and swift, while the merge area exit flow q_{out} is correspondingly low. As the arriving flow, and hence N, increase, merging conflicts may increase, but q_{out} increases as well until, for a specific critical value N_{cr}, the exit flow reaches the downstream capacity q_{cap}. If N increases beyond N_{cr}, merging conflicts become more serious, leading to substantial vehicle decelerations and eventual accelerations that reduce the exit flow to lower values q_c, where $q_{cap} - q_c$ is the capacity drop due to congestion, which is deemed to depend on vehicle acceleration at the congestion head. Under these conditions, real-time control of the arriving flow may be employed in order to maintain the number of vehicles N in the merge area close to its critical value N_{cr}. This is similar to local ramp metering measures (Papageorgiou and Kotsialos, 2002) where, in contrast, only a part of the arriving traffic flow (i.e. only the on-ramp flow) is controlled so as to maximize the merge area throughput see also Papageorgiou et al. (2008) for a more comprehensive description.

2.2. Control devices and real-time measurements

Merging traffic control could be applied by use of different control devices aiming at a smooth, safe and efficient merging of vehicles. A possible control device to regulate the arriving flow at work zone areas are traffic lights. A significant issue when applying work zone traffic control, which had been neglected in previous works but is crucial for a successful implementation, is the proper positioning of the traffic lights so as to achieve the most efficient merging of vehicles. More specifically, the traffic lights should be placed sufficiently upstream from the merge area, so that the vehicles, starting from the traffic lights at low speed, have enough time to accelerate and reach the appropriate speed for orderly and efficient merging, i.e. a speed that corresponds roughly to capacity flow in Fig. 1b (critical speed). This implementation of a mainstream metering policy will of course lead to queue formation at the traffic lights, but the achieved outflow at the merge area will be higher than the outflow under congestion (due to mitigation of the capacity drop), thus resulting in a corresponding reduction of the total time spent in the network. Traffic lights can be applied to all lanes simultaneously or to individual lanes separately. If necessary, traffic lights may be installed only on a part of the motorway lanes, while traffic flow on other lanes is allowed to enter the merge area freely (e.g. by-pass lanes for HOV, buses, emergency vehicles).

In order to apply feedback control and maintain the number of vehicles N close to N_{cr}, real-time measurements or estimates of N are needed. A frequently practiced way of estimating N is by use of ordinary loop detectors placed at appropriate positions (Vigos et al., 2008). Alternatively, one may employ occupancy measurements and target a critical occupancy value o_{cr} (instead of N_{cr}) as in ALINEA ramp metering.

2.3. Control algorithm

The outlined merge area and control device characteristics lead to the following process to be controlled:

- The control input is the entering flow $q(k)$ to be released from the traffic lights upstream of the merge area.
- This flow reaches, after a small time delay (corresponding to the vehicle travel time), the merge area.
- From the conservation law, the number of vehicles $N(k)$ in the merge area results from a balance of the (time-delayed) inflow and the outflow q_{out}.
- The outflow q_{out} depends on the number of vehicles $N(k)$ according to Fig. 1b.

In summary, the process under control is a SISO (single-input–single-output) time-delayed nonlinear first-order dynamic system with control input $q(k)$ and output $N(k)$. It is well-known (see e.g. (Seborg et al., 1989)) that such a system may be efficiently and robustly controlled by use of integral (I-type) or, in case of more substantial time delays, of proportional–integral (PI-type) controllers.

The control algorithm makes use of real-time measurements or estimates of the number of vehicles N or occupancy o in the network in order to maintain $N \approx N_{cr}$ or $o \approx o_{cr}$, which maximizes the merge area exit flow. A crucial investigation in this study concerns the appropriate positioning of the traffic lights. Since different distances of the traffic lights from the merge area give rise to correspondingly different (partly substantial) time-delays, the feedback algorithm employed in this study for merging traffic control is an extension of the local ramp metering strategy ALINEA (Papageorgiou et al., 1991, 1997). ALINEA is an integral (I-type) feedback regulator given by the equation

$$q(k) = q(k-1) + K_R[\hat{o} - o(k-1)] \tag{1}$$

where $k = 1, 2, \ldots$ is the discrete-time index, $q(k)$ denotes the controlled entering flow to be implemented during the next period k, $K_R > 0$ is a regulator parameter, $o(k-1)$ is the last measured occupancy (%) and $\hat{o}$ is a set (desired) value for the downstream occupancy of the motorway. A typical set-value is $\hat{o} = o_{cr}$, in which case the motorway exit flow becomes close to q_{cap}. The same equation can be used if the number of vehicles N is measured, instead of the occupancy percentage. Within the microscopic simulation environment, the performance of this regulator was sometimes sluggish, possibly due to the involved time delays; according to the closed-loop analysis undertaken in Papageorgiou et al. (1991), the closed-loop dynamic behavior may be rendered more tight if a proportional term is added to the regulator (1). Moreover, the addition of a proportional term was found in Wang and Papageorgiou (2006) to lead to low sensitivity with respect to varying distances between the traffic lights and the bottleneck area, which is a highly desired feature in the present context. Thus, a

proportional–integral (PI-type) regulator was finally employed as better suitable for the present investigation. The PI-type regulator reads (now based on the number of vehicles N rather than on occupancy)

$$q(k) = q(k-1) - K_P[N(k) - N(k-1)] + K_I[\widehat{N} - N(k)] \tag{2}$$

where K_P and K_I denote the regulator parameters for the proportional and integral terms, respectively, that must be suitably specified, and $\widehat{N}$ is a set (desired) value for the downstream number of vehicles. The flow calculated with (2) is truncated if it exceeds a range $[q_{min}, q_{max}]$ of pre-selected admissible values so as to avoid the wind-up phenomenon.

The control algorithm is activated at each time interval T (in s) and calculates the entering flow $q(k)$ (in veh/h) to be implemented in the next interval k via appropriate operation of the control devices (traffic lights). There are different possible metering policies to translate the decision of the control strategy (i.e. the flow $q(k)$) into corresponding traffic light settings; e.g. one-car-per-green, n-cars-per-green, full traffic cycle, discrete release rates, etc. (see Papageorgiou and Papamichail, 2008). A full traffic cycle policy is employed here, so as to maximize the resulting flow capacity of the traffic lights. The flow to be implemented in the next control period T may be distributed equally among the motorway lanes via corresponding individual traffic lights for each lane; while a shift (offset) should be applied for the signal cycle start of each traffic light relative to the cycles of the other traffic lights, so as to enable (to the extent possible) a continuous flow and avoid simultaneous vehicle departures from all lanes (or no departures during red).

3. Application setup

The described real-time work zone merging control concept is implemented, via microscopic simulation, at a hypothetical work zone infrastructure featuring 3 arriving lanes and 2 exiting lanes as depicted in Fig. 2. Clearly, there may be many different work zone geometries in real applications, but this should not alter the essence of the control approach, of its properties and of the achievable improvements. The total length of the simulated motorway stretch is 5 km (to accommodate any forming queue length), while the trapezoidal merging area, which is situated 100 m before the end of the motorway stretch, is 50 m long. A speed limit of 80 km/h is applied along the whole motorway stretch. The capacity q_{cap} of the motorway upstream of the work zone area is sufficiently high to accommodate the investigated demand scenario, while the downstream capacity is reduced due to the lane drop and was found empirically to amount to 4800 veh/h (for a traffic flow including 20% trucks). Another feature of the described infrastructure is that the left-most lane of the motorway, which is the high-speed lane, is reserved only for cars, while trucks are allowed to use only the other two lanes, as in several real motorways (see Fig. 2). For the collection of measurements, for operation or evaluation, detectors have been placed at different positions along the stretch, as displayed in Fig. 2.

The control concept was implemented for a representative demand scenario, which has a duration of 2 h and trapezoidal profile, see Fig. 3. At the beginning of the simulation, the average demand at the motorway entrance starts at a low value (50% of the highest traffic demand). The demand increases gradually within the first 30 min, until it reaches a peak demand of 5400 veh/h; and remains at this value for the next 30 min. During this time period the traffic demand exceeds the merge area capacity q_{cap} which is expected to lead to congestion and reduced efficiency of the infrastructure. During the time period $t \in [60\,\text{min}, 90\,\text{min}]$, the demand reduces gradually back to the initial low value (50% of the highest demand) and remains at that value for the rest 30 min, i.e. until the end of the simulation. Any queues must be resolved at the end for all scenarios, to obtain comparable results.

The vehicle types included in the demand scenario are cars and trucks. The trucks represent an average of 20% of the total traffic demand, and this percentage remains constant (in average) throughout the simulation. While determining N for the control algorithm, trucks are counted as equivalent to two cars. Since the same equivalence is used for the regulator setpoint $\widehat{N}$, the system is robust to this assumption.

The regulator (Eq. (2)) is activated every $T = 30$ s and receives the real-time measurement of the number of vehicles N included in the merge area (calculated from detectors 1 and 3 in Fig. 2, which comprises also a short upstream motorway part to better capture possible early vehicle merging) to calculate the entering flow $q(k)$ to be implemented in the next control period k so as to maintain $N \approx N_{cr}$. The new entering flow to be implemented is not allowed to exceed the range $q \in [4000, 6000]$ veh/h, i.e. a minimum and maximum flow, respectively. Specifically, the minimum admissible flow q_{min} was selected lower than the downstream capacity to enable a sufficient margin for regulator action; for the same reason the maximum admissible flow q_{max} was selected sufficiently large and higher than q_{cap}. The traffic cycle is fixed and equal

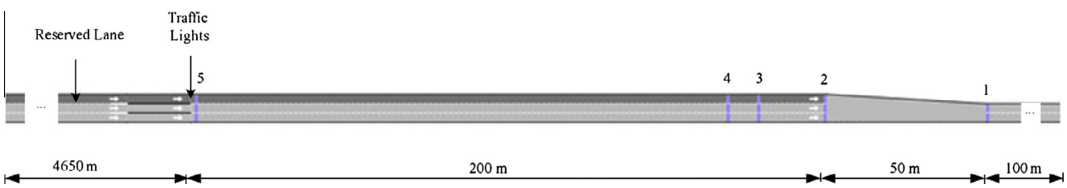

Fig. 2. The work zone infrastructure under consideration. (Numbers correspond to emulated detector locations.)

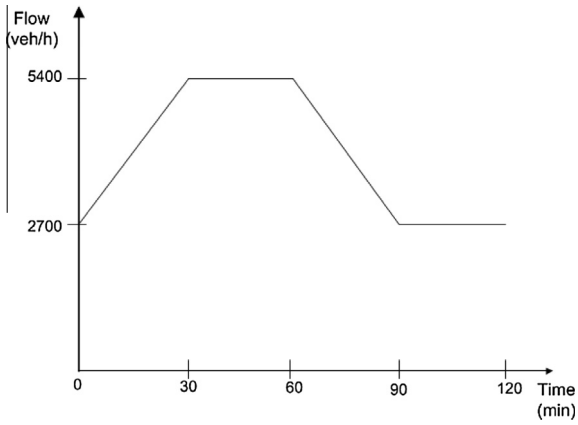

Fig. 3. Demand scenario.

to the control interval (30 s); while the green and red phases are calculated appropriately to implement the ordered flow $q(k)$, with a minimum red phase of 3 s being considered for safety reasons. More specifically, the flow q (veh/h) is translated into a corresponding green phase G (in s) via the following equation

$$G = (q \cdot T)/n \cdot S \tag{3}$$

with subsequent application of constraints for the green phase G mentioned earlier. In (3), S (in veh/h) is the saturation flow per lane and n is the number of lanes, which in this application is equal to 3. Thus, the minimum-red constraint of 3 s leads to a maximum green $G_{\max} = 27$ s, while, for the given $q_{\min}$ and a saturation flow of 2000 veh/h, the minimum green resulting from (3), is 20 s. The same green phase is implemented at all motorway lanes, albeit with an offset of the cycle start as mentioned earlier. In view of the minimum-red constraint, the maximum implementable flow resulting from (3) is 5400 veh/h. Note that the virtually inevitable inaccuracies resulting from (3) do not affect the efficiency and accuracy of the control action, thanks to the feedback character of the regulator which is continuously striving to compensate these and many other kinds of errors; see also the related analysis in Papageorgiou et al. (1991) and Kotsialos et al. (2006).

The specification of appropriate regulator parameter values was conducted manually, via trial-and-error. Specifically, various sets of values were tested through a series of simulation runs considering a specific position for the traffic lights. As mentioned earlier, the employed regulator is not very sensitive to the distance between the measurement point and the control device; in fact, the parameter values resulted from this investigation were found to work equally well for different positions of the traffic lights.

In order to specify a preliminary but reasonable position for the traffic lights, a primary investigation was conducted, using a fixed flow rate for the traffic lights and setting the traffic lights at different locations upstream of the work zone area, from 50 m up to 300 m in steps of 50 m. The distance for which vehicles had enough time to acquire a speed close to the critical value, before reaching the merge area, was around 200 m. Therefore, the regulator fine-tuning experiments were conducted by positioning the traffic lights 200 m upstream from the merge area (Fig. 2).

The described infrastructure was simulated by use of the microscopic simulator AIMSUN v.6.0.6 (TSS, 2009), using the simulator's default parameters and a simulation time step of 0.1 s. The implementation of the control strategy was done via the AIMSUN API (Application Programming Interface), which allows the user to emulate a real-time control environment. Specifically, the simulator delivers at every control period T the number of vehicles N between detectors 1 and 3; based on these measurements, the control software calculates the corresponding traffic light settings and returns them to the microsimulator for application. Since the AIMSUN simulator model is stochastic, different replications with different random seeds may produce different results. For this reason, 10 replications with different random seeds were carried out for each examined scenario.

Microscopic simulators are known to feature limited accuracy when modeling merging situations; therefore, the quantitative results obtained in the present study might not materialize at the exact same level in practice. However, more realistic merging behavior is not expected to have a structural impact on the behavior of the process under control; it may merely lead to different model parameters (e.g. of Fig. 1b) and, possibly, to slightly different traffic light distance for efficient control. The important message, highlighted and demonstrated via the provided simulations, is that the position of the traffic lights should allow for adequate vehicle merging speed. Also, the structure of the regulator will remain adequate in practical applications, and only a re-tuning of its parameters may be necessary. It should be emphasized that the regulator does not use any traffic flow model; it merely reacts to real-time measurements, both in the present simulations and in a potential field application. Thus, given the known inherent robustness of feedback regulators, a similar kind of control behavior (and resulting improvements) can be expected in practice.

4. Simulation results

4.1. No-control case

In the no-control case, arriving vehicles enter the merge area and exit without any serious problem, as long as the arriving demand is sufficiently low. When the demand increases (peak period) beyond the work zone capacity, vehicle merging conflicts are observed that lead to vehicle decelerations and formation of congestion. Congestion spills back several kilometers, but without ever reaching the simulated network entrance. The mean of the resulting average vehicle delays (AVD) (in s/veh/km) for 10 replications is 38.1 while the minimum and maximum values are 24.7 and 51.7, respectively.

The trajectories in Fig. 4a and b display the merge area outflow q_{out} (collected from detector 1) and the number of vehicles in the merge area, respectively, for one particular simulation run with $AVD = 37.97$ s/veh/km, which is closest to the mean AVD of the 10 replications. Until about $t = 40$ min, the number of vehicles N in the merge area is slowly increasing (as a consequence of the increasing demand), while the merge area outflow is seen to follow the increase of arriving demand, reaching approximately 5000 veh/h in average. At $t = 40$ min, the number of vehicles in the merge area increases steeply due to serious merging conflicts that lead to a speed breakdown, and this congested traffic situation becomes stationary until $t = 110$ min. The outflow during this time period is reduced to around 4150 veh/h in average due to the merge area congestion (capacity drop). Fig. 5a shows a snapshot of the simulated stretch at $t = 60$ min with the formed congestion. After $t = 110$ min, when the queue dissolves, the number of vehicles in the merge area is seen to drop, and the outflow reduces to lower values due to the decreased demand. Vehicle speed measurements collected upstream of the merge area, at detector 4, indicate that, during the peak period queuing, there is a serious speed drop down to around 20 km/h in average, while in the rest of the simulation horizon the average vehicle speed is around 82 km/h in average.

4.2. Merging control case

When real-time merging traffic control is applied, the maximum admissible flow $q_{max} = 6000$ veh/h is ordered by the regulator for as long as the number of vehicles N in the merge area is lower than the set value $\widehat{N}$ in PI-ALINEA's Eq. (2). As the demand increases, N increases as well, and when $N(k)$ approaches $\widehat{N}$, the controller starts its actual operation aiming at maintaining $N(k)$ close to $\widehat{N}$. At this time, a queue is formed upstream of the traffic lights (since the arriving demand is higher than the work zone capacity) which propagates backwards, but without reaching the entrance of the simulated system. Fig. 5b shows a snapshot of the simulated stretch at $t = 60$ min with the formed queue upstream of the traffic lights.

A critical question at this point is the proper choice of the regulator parameters K_P, K_I and of the set-point $\widehat{N}$. The regulator parameters may be manually fine-tuned following some practical trial-and error rules from Control Engineering (see e.g. Seborg et al., 1989); once appropriate values have been found, the regulator is known to be little sensitive to related

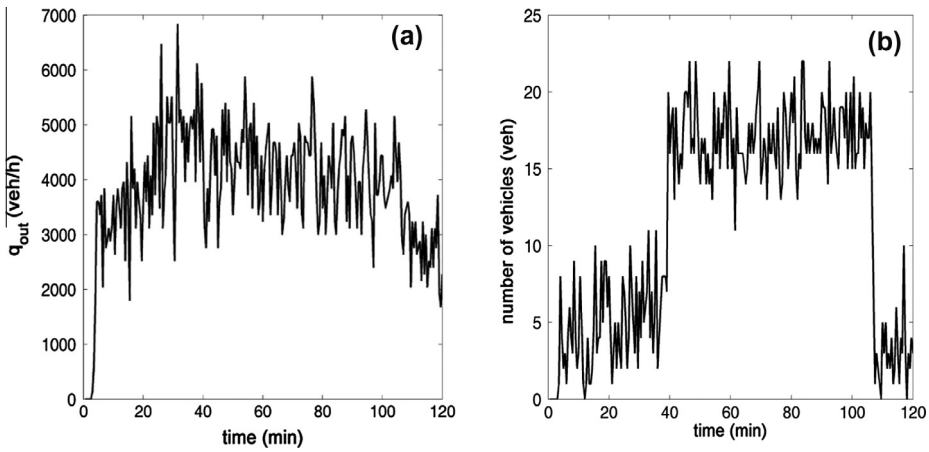

Fig. 4. (a) Outflow q_{out} and (b) number of vehicles N in the no-control case.

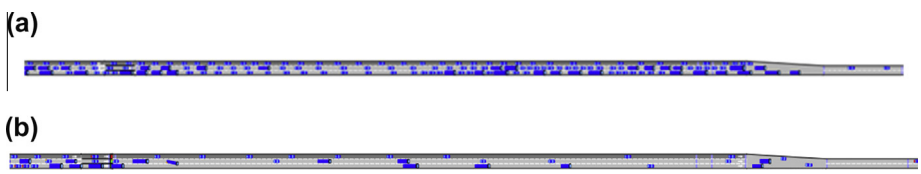

Fig. 5. Screenshot of simulation ($t = 60$): (a) without control and (b) with merging traffic control.

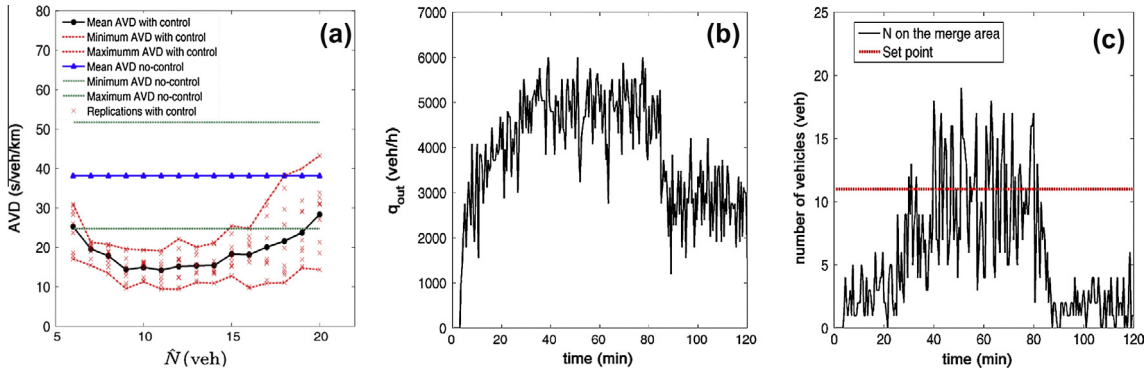

Fig. 6. (a) Average vehicle delay versus $\widehat{N}$ values with and without control; (b) outflow q_{out} and; (c) number of vehicles N for the control case.

moderate variations. Of course, more substantial changes in the values of the regulator parameters may have a sensible impact on the closed-loop behavior as demonstrated in the next section. The manually derived values of regulator parameters K_P and K_I are 150 h^{-1} and 6 h^{-1}, respectively. The value of $\widehat{N}$, on the other hand, should be selected such that the work zone throughput is maximized, according to Fig. 1b. In a field investigation, this may be achieved by gradually incrementing $\widehat{N}$ and monitoring the measured outflow, until a maximum throughput is obtained. In the current case, the investigation of the $\widehat{N}$ value is carried out through a series of simulation experiments with different (integer) $\widehat{N}$ values within the range $\widehat{N} \in [6, 20]$ veh. For each investigated $\widehat{N}$ value, the mean AVD of 10 replications is obtained. Fig. 6a displays, for every investigated $\widehat{N}$ value, the corresponding AVD values for the 10 replications as well as the mean, minimum and maximum AVD of all replications. The mean, minimum and maximum AVD values for 10 replications of the no-control case are also displayed on the same figure for comparison. According to the displayed results, the mean AVD value is minimized in the range of $\widehat{N} \in [9, 14]$ veh, and, particularly for $\widehat{N}$ equal to 11 veh, it takes the lowest value, which corresponds to the critical value mentioned earlier. For lower $\widehat{N}$ values, the system operates at undercritical conditions (Fig. 1b) and the merge area "starves for flow"; while, for higher set points, the merging conflicts are increasingly frequent and serious, leading to reduced throughput. It is also noteworthy that the mean AVD for all investigated $\widehat{N}$ is significantly lower than the corresponding value of the no-control scenario. Particularly, for $\widehat{N} = 11$ veh, the mean AVD is 14.25 s/veh/h, which is 63% lower than the mean AVD of the no-control case.

Fig. 6b and c display the merge area outflow q_{out} and the number of vehicles in the merge area, respectively, for one particular replication with $\widehat{N} = 11$ veh and $AVD = 13.83$ s/veh/km, which is closest to the mean AVD value of the corresponding 10 replications. The number of vehicles in the merge area is maintained around the set-point $\widehat{N} = 11$ veh (red dashed line in the figure) during the peak period. The observed spikes are due to stochastic arrivals, but also due to some occasional vehicle merging conflicts that may occur and lead to vehicle decelerations in the merge area; the appropriate reaction of the regulator in such cases, guarantees that the number of vehicles in the merge area remains around the set-point in average. The outflow q_{out} maintains its average value around 4850 veh/h during the peak period (between $t = 30$ min and $t = 60$ min), and beyond the peak period for some 20 more minutes because of the queued vehicles at the traffic lights. This marks a reduction of the congested period by 30 min (or around 43%) compared to the no-control case. The mean vehicle speed upstream of the merge area is significantly increased during the peak hours to 70 km/h, except for the occasional departures to lower values due to corresponding merging conflicts (see Fig. 7b).

4.3. Investigation of the position of the traffic lights

What is a good distance of the traffic light location, measured from the upstream border of the merging area? For efficient (capacity-flow) merging, vehicles should reach the merge area at a (quasi-stationary) speed that corresponds roughly to the critical speed of Fig. 1b (typically some 70 km/h). To this end, the location of the traffic lights should be sufficiently upstream of the merge area in order to allow for the vehicles to accelerate, reach the critical speed and pass through the merge area efficiently, i.e. without strong acceleration and the related gaps forming in front of slowly accelerating vehicles (e.g. trucks). On the other hand, the traffic light distance should not be unnecessarily long, so as to limit the space required for the control concept implementation and the upstream reach of the traffic light queue.

To investigate this issue, a range of possible distances from 30 up to 400 m, were considered, and 10 replications were simulated for each of them. The obtained mean AVD values as well as the acquired vehicle speed when approaching the merge area (save the occasional drops due to short-lasting merging conflicts) are the main evaluation criteria for the determination of the optimum position of the traffic lights. Fig. 7a depicts, for every investigated traffic lights position, the resulting AVD values for the 10 replications, as well as the mean, minimum and maximum AVD of all replications. The mean, minimum and maximum AVD values of the 10 replications of the no-control case are also displayed on the same figure for comparison. The mean average vehicle delay is seen to be low and virtually constant when the traffic lights are located

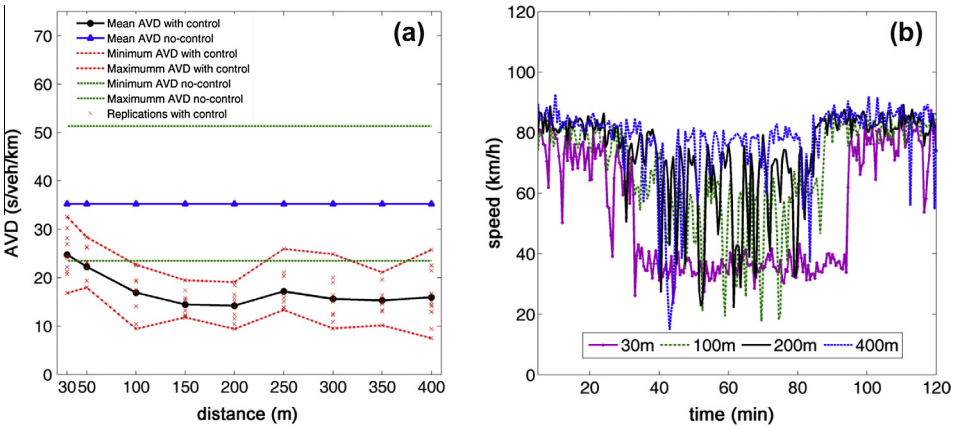

Fig. 7. (a) Average vehicle delay and (b) speed versus time for different traffic lights positions.

150 m upstream of the merge area or more. In contrast, when traffic lights are placed closer to the merge area, higher AVD values are seen to result.

Fig. 7b displays the trajectories of the speed measurements collected just upstream of the merge area (detector 4 in Fig. 2), only for selected traffic lights positions, for readability. When the traffic lights are placed very close to the merge area, e.g. at 30 m or 50 m, vehicles do not have sufficient space to accelerate, and indeed it can be observed that the mean speed value during the peak period is quite low, i.e. lower than 45 km/h. For a distance equal to 100 m, the merging vehicle speed increases, reaching 60 km/h in average. For the traffic lights position of 200 m upstream of the merge area, the achieved merging vehicle speed has increased to a mean value around 70 km/h. For even longer distances, e.g. 400 m, the speed is even more increased to around 80 km/h. The observed occasional speed drops are due to temporal sharp vehicle conflicts. Apparently, distances less than 150 m are less appropriate, and, particularly for distances less than 50 m, the system performance comes quite close to the no-control case, because the capacity drop is only partially avoided. For distances more than 200 m, there is no further improvement, since the critical merging speed has been reached. Thus, 200 m is the most appropriate distance, as it is preferable to have the traffic lights closer to the merge area.

Clearly, these quantitative results may not be accurately transferable to different work zone types, different geometrical characteristics, and, last not least, to the real traffic and merging conditions. For example, a traffic lights distance of, say, 250 m, may prove more appropriate for a specific real work zone setting. However, the basic underlying relations between traffic lights distance, the achievable vehicle merging speed and the resulting flow efficiency, as investigated and demonstrated in this study, are expected to hold for a variety of infrastructure types and traffic conditions in practice.

5. Application of AFT to the merging control strategy

Recently, a learning/adaptive algorithm called AFT (Adaptive Fine-Tuning) was proposed (Kosmatopoulos, 2009; Kouvelas, 2011; Kouvelas et al., 2011) to enable automatic fine-tuning of traffic control systems (TCS), so as to reach a better, in some cases even the best, measurable performance with the applied control strategy. In this section, the automatic fine-tuning method is employed for the motorway work-zone merging control concept in order to test and demonstrate its capabilities while automatically fine-tuning the regulator parameters of the PI-ALINEA control strategy under different conditions.

The performance J of any traffic control system over a period of time (e.g. one day) depends on two factors: (a) external inputs (or disturbances) d; and (b) control strategy parameters θ. The AFT algorithm attempts to approximate the performance function $J(d,\theta)$ by use of real measurements. More specifically, the original algorithm employs a polynomial-like approximator (similar to a neural network) that approximates, based exclusively on available measurements, the unknown nonlinear performance function of the problem. For the present application, a polynomial approximator as well as a Support Vector Machine (SVM) model (see Burges, 1998) are used. In fact, an SVM model was recently used by Giannakis et al. (2011) within the AFT frame in order to fine-tune the parameters of a building's controller, with excellent results. The performance function (to be minimized) for the present problem is the average vehicle delay (AVD). The AFT algorithm (Fig. 8) runs iteratively, where by each iteration corresponds to the duration of the demand scenario (2 h), or of one day in field applications.

Fig. 8 illustrates the working principle of AFT algorithm. The basic procedure of the self-tuning process may be summarized as follows:

- The traffic flow process (here, work zone traffic) is controlled in real time by a control strategy (here, the PI-regulator) which includes a number of parameters.

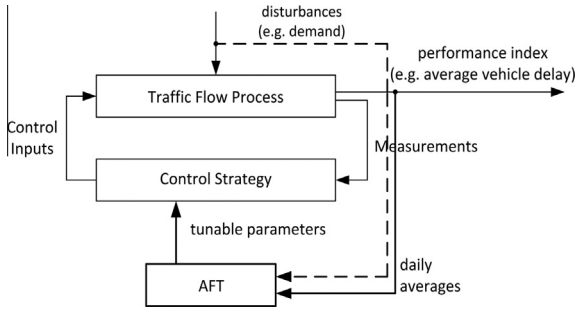

Fig. 8. Working principle of AFT for automatic calibration.

- At the end of appropriately defined periods (e.g. at the end of each day; here, at the end of each simulation run), AFT algorithm receives the values of the real (measured) performance index (here, AVD) and of the external inputs (here, the demand).
- Using the measured performance (the samples of which increase iteration by iteration), AFT algorithm updates and enhances the approximation of the function $J(d, \theta)$.
- Based on the approximate performance function, AFT calculates new parameter values to be applied at the next period (here, in the subsequent simulation run) in an attempt to improve the system performance.
- This (iterative) procedure is continued over many periods until a (local) maximum in performance is reached.

The AFT algorithm is started with some initial values for the parameters to be tuned. When using an SVM approximator, the algorithm convergences fast to a close local minimum of the performance function and remains there; while the original AFT (using the polynomial approximator) may feature more significant "jumps" in the parameter space, with correspondingly stronger fluctuations of the performance function. In order to use SVM, a sufficient initial set of training data is needed to be available for fitting. Therefore, the original AFT is applied for the first 10 iterations, before switching to the SVM usage. In particular, for the first 10 iterations, AFT explores a wide region of alternative sets of parameters; following which the SVM is used and the algorithm converges to a close local minimum without exploring other feasible regions.

Due to the non-convex nature of the investigated problem, a global minimum (i.e. unique regulator parameter values) cannot be guaranteed, hence the algorithm may converge to different local minima depending on different starting points. Thus, the AFT algorithm is applied in the following for three control scenarios with different initial values for the regulator parameters in order to investigate the behavior of the algorithm under different conditions.

In a first experiment, the AFT algorithm is applied with the initial values of the tunable parameter set as following: $K_P = 150\ \mathrm{h}^{-1}$, $K_I = 6\ \mathrm{h}^{-1}$ and $\widehat{N} = 11$ veh. These are the optimized parameter values derived via the trial-and-error procedure, which means that these values represent already a "good" starting point for the AFT algorithm; and it is interesting to investigate whether AFT can further improve the control performance. Fig. 9 displays the AVD trajectories, for each of the three control scenarios, delivered for one run of the AFT algorithm. For the first control scenario, some strong fluctuations

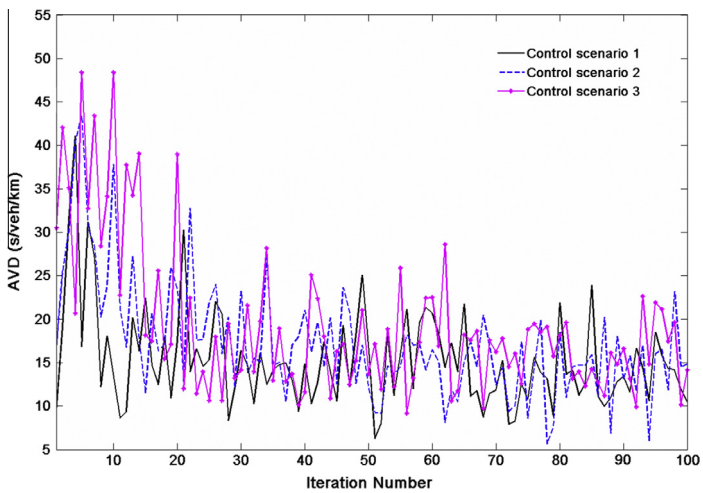

Fig. 9. Average vehicle delay for control scenarios 1, 2 and 3.

are observed in the first 20 iterations, but during the rest of the fine-tuning period the AFT algorithm achieves keeping the AVD at lower values in average, whereby the continuing smaller AVD fluctuations are mainly due to the stochastic simulation results. The AFT algorithm converges to the following values for the parameters: $K_P = 81\ h^{-1}$, $K_I = 3\ h^{-1}$ and $\widehat{N} = 8.4$ veh. In Fig. 10a–c, the trajectories of each tunable parameter are also displayed. During the first iterations, the observed oscillations are strong, as the AFT algorithm is learning the system's behavior by experimenting with different sets of parameters. However, after a few iterations the magnitude of the parameter oscillations is decreasing and, eventually, the convergence to a local minimum is clearly visible.

The delivered set of values was then applied within the work-zone control concept in order to compare the system performance before and after the use of the AFT algorithm. The resulting mean AVD for 10 replications is 13.0 s/veh/km, which is around 9% lower than the mean AVD of the initial control scenario before using the AFT algorithm. This result indicates that AFT can improve the control performance resulting from a careful manual fine-tuning.

In a second AFT experiment, the starting values for the tunable parameters were set $K_P = 50\ h^{-1}$, $K_I = 10\ h^{-1}$ and $\widehat{N} = 10$ veh. This set of parameters was arbitrarily selected in order to investigate the algorithm's behavior in case of a reasonable, but not really good starting point. The simulated 10-replication results for the control scenario using these parameter values lead to quite high AVD values, with a mean AVD around 21 s/veh/km. Fig. 9 presents the corresponding trajectory of the AVD values during the AFT fine-tuning process. It is visible that the obtained AVD values exhibit perturbations due to the search process of the AFT algorithm, but finally AFT achieves to locate a satisfactory set of parameters that leads to low AVD values. The corresponding fluctuations of the tunable parameter trajectories in Fig. 10a–c are initially quite strong, but after a while the AFT algorithm is seen to converge to specific values, which reflects an efficient fine-tuning of the regulator parameters. Specifically, the AFT algorithm converges to $K_P = 74\ h^{-1}$, $K_I = 4.5\ h^{-1}$ and $\widehat{N} = 8.6$ veh, and the corresponding mean AVD for 10 replications is 14.09 s/veh/km, which is around 34% lower than the mean AVD of the starting control scenario before using AFT. The performance with the obtained parameter values is even slightly better than the manually fine-tuned regulator parameters. This result indicates that AFT may lead to strong control performance improvement if no manual fine-tuning was carried out.

In a third AFT experiment, the starting values for the tunable parameters were $K_P = 200\ h^{-1}$, $K_I = 50\ h^{-1}$ and $\widehat{N} = 50$ veh with a corresponding mean AVD, for 10 replications, around 36 s/veh/km. These values correspond to a very "bad" set of regulator parameters and they are selected in order to investigate the performance of the AFT algorithm and the values that it is going to converge to, under very unfavorable starting conditions. The corresponding trajectory of the AVD values during the fine-tuning process is presented in Fig. 9. During the first iterations, the AVD values are quite high; however, after the 20th iteration, the AFT algorithm manages to lead and keep the AVD at lower values. The AFT algorithm converges to $K_P = 204.18\ h^{-1}$, $K_I = 31.32\ h^{-1}$ and $\widehat{N} = 12.13$ veh. The obtained values of K_P and K_I are relatively close to the initial ones, but the value $\widehat{N}$ has been reduced a lot. The resulting mean AVD for 10 replications is 15 s/veh/km which is 57% lower than the mean AVD of the starting control scenario, albeit slightly worse than in the manually fine-tuned case.

Table 1 summarizes the obtained results of the AFT experiments. The experiments demonstrate that the AFT algorithm can improve the system performance independently of the starting points (except in the rare case where the starting point is a local minimum). Nevertheless, the selection of, at least roughly, appropriate starting values for the regulator parameters

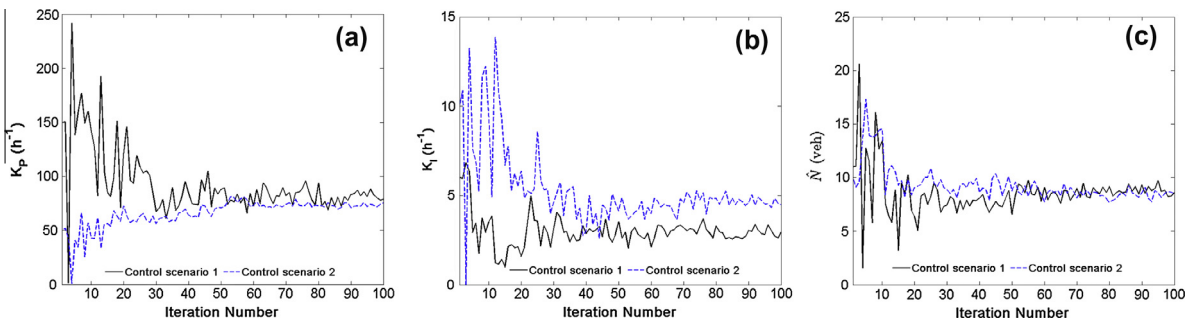

Fig. 10. Parameter tuning over iterations for experiments 1 (continuous line) and 2 (discontinuous line): (a) K_P parameter; (b) K_I parameter; (c) set-point $\widehat{N}$.

Table 1
Summarized results of the AFT algorithm experiments.

Experiment	Starting parameter values				Final parameter values				AVD improvement (%)
	K_P (h^{-1})	K_I (h^{-1})	$\widehat{N}$ (veh)	AVD (s/veh/km)	K_P (h^{-1})	K_I (h^{-1})	$\widehat{N}$ (veh)	AVD (s/veh/km)	
1	150	6	11	14.25	81	3	8.4	13.0	9
2	50	50	10	21.0	74	4.5	8.6	14.09	34
3	200	50	50	36.0	204.18	31.32	12.13	15.0	57

(e.g. derived from manual fine-tuning) may be necessary in order to achieve best performance of the utilized control strategy.

6. Conclusions

A previously proposed control scheme for real-time merging traffic control at work zones with lane drop was farther developed, enhanced and demonstrated for a hypothetical work zone motorway infrastructure within a microscopic simulation environment. The control algorithm used for work zone management is an extension of the well-known local ramp metering strategy ALINEA, while the control devices to implement the control algorithm decisions are traffic lights located sufficiently upstream of the work zone area. The reported research addressed the appropriate distance between the traffic lights and the merge area, and demonstrated its significance for throughput maximization (or equivalently delay minimization) via avoidance of the capacity drop. A further investigation in this research is related to the fine-tuning procedure needed for the calibration of the control algorithm parameters. In particular, after the manually conducted fine-tuning procedure, the recently proposed learning/adaptive algorithm AFT was applied in order to seek for better regulator parameter values which lead to improved performance of the utilized control strategy. The significance of this investigation is that the AFT algorithm can be applied similarly in the field, based on real measurements, to optimize the control system performance.

References

Burges, C.J.C., 1998. A Tutorial on support vector machines for pattern recognition. Data Min. Knowl. Disc. 2, 121–167.
Carlson, R.C., Papamichail, I., Papageorgiou, M., Messmer, A., 2010. Optimal mainstream traffic flow control involving variable speed limits and ramp metering. Transp. Sci. 18, 193–210.
Carlson, R.C., Papamichail, I., Papageorgiou, M., 2011. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. IEEE Trans. Intell. Transp. Syst. 12, 1261–1276.
FHWA, 2005. Developing and Implementing Transportation Management Plans for Work Zones: Report of the Federal Highway Administration. Department of Transportation, US.
Gazis, D.C., Foote, R.S., 1969. Surveillance and control of tunnel traffic by an on-line digital computer. Transp. Sci. 3 (3), 255–275.
Giannakis, G.I., Kontes, G.D., Kosmatopoulos, E.B., Rovas, D.V., 2011. A model-assisted adaptive controller fine-tuning methodology for efficient energy use in buildings. 19th IEEE Mediterranean Conference on Control and Automation, Corfu, Greece, June, pp. 49–54.
Kosmatopoulos, E.B., 2009. An adaptive optimization scheme with satisfactory transient performance. Automatica 45 (3), 716–723.
Kouvelas, A., 2011. Adaptive Fine-Tuning for Large-Scale Nonlinear Traffic Control Systems. Ph.D thesis. Technical University of Crete, Chania, Greece.
Kouvelas, A., Aboudolas, K., Kosmatopoulos, E.B., Papagorgiou, M., 2011. Adaptive performance optimization for large-scale traffic control systems. IEEE Trans. Intell. Transp. Syst. 12 (4), 1434–1445.
Kotsialos, A., Papageorgiou, M., Hayden, J., Higginson, R., McCabe, K., Rayman, N., 2006. Discrete release rate impact on ramp metering performance. IEE Proc. Intell. Transp. Syst. 153, 85–96.
Lentzakis, A.F., Spiliopoulou, A.D., Papamichail, I., Papageorgiou, M., 2008. Real-time work zone management for throughput maximization. 87th Transportation Research Board Annual Meeting, Washington, DC, USA.
Lin, P.W., Kang, K.P., Chang, G.L., 2004. Exploring the effectiveness of variable speed limit controls on highway work-zone operations. J. Intell. Transp. Syst. 8, 155–168.
Murashige, Y., 2011. Countermeasure strategies against traffic congestion on motorways in Japan. Proc. Social Behav. Sci. 16 (110–119), 2011.
Papageorgiou, M., Haj-Salem, H., Blosseville, J.-M., 1991. ALINEA: a local feedback control law for on-ramp metering. Transp. Res. Rec. 1320, 58–64.
Papageorgiou, M., Haj-Salem, H., Middelham, F., 1997. ALINEA local ramp metering: summary of field results. Transp. Res. Rec. 1603, 90–98.
Papageorgiou, M., Papamichail, I., 2008. Overview of traffic signal operation policies for ramp metering. Transp. Res. Rec. 2047, 28–36.
Papageorgiou, M., Papamichail, I., Spiliopoulou, A.D., Lentzakis, A.F., 2008. Real-time merging traffic control with applications to toll plaza and work zone management. Transp. Res. Part C 16, 535–553.
Papageorgiou, M., Kotsialos, A., 2002. Freeway ramp metering: an overview. IEEE Trans. Intell. Transp. Syst. 3, 271–281.
Seborg, D., Edgar, T.F., Mellichamp, D.A., 1989. Process Dynamics and Control. Wiley, New York.
TSS – Transport Simulation Systems, 2009. AIMSUN Users Manual Version 6, Barcelona, Spain.
Vigos, G., Papageorgiou, M., Wang, Y., 2008. Real-time estimation of vehicle-count within signalized links. Transp. Res. Part C 16, 18–35.
Wang, Y., Papageorgiou, M., 2006. Local ramp metering in the case of distant downstream bottlenecks. IEEE Intelligent Transportation Systems Conference, Toronto, Canada, September 17–20, pp. 426–431.
Wei, H., Pavithran, M., 2006. Concept of dynamic merge metering approach for work zone traffic control. 11th IFAC Symposium on Control in Transportation Systems. Delft, The Netherlands, August 29–31, pp. 374–379.