Dynamic speed limit control to resolve shock waves on freeways – Field test results of the SPECIALIST algorithm

A. Hegyi, S.P. Hoogendoorn

Abstract—We present the real-world test of the SPECIALIST algorithm in which dynamic speed limits were used to resolve shock waves on freeways. The real-world test was performed in the period September 2009–February 2010 on a 14 km long stretch on the Dutch A12 freeway.

For the evaluation of the algorithm various performance measures were determined for each activation of the speed limits. The results show that the SPECIALIST algorithm can resolve shock waves in nearly 80% of the cases when the algorithm activated for shock waves. However, in approximately 50% of the activations the algorithm activated for jam types other than shock waves, in which case the effectivity was 40–50%. Due to the tuning the stability of the traffic flow could be significantly improved, and the number of average activations per day was increased from 1.5 to 2.9 activations per day.

The SPECIALIST algorithm was not only evaluated for traffic performance, but also for the correctness of the algorithm for real traffic in terms of expected qualitative behavior. Although the general operation of the algorithm is according to the theoretical expectations, some points for further improvements are identified during the test.

I. INTRODUCTION

In earlier publications we presented an approach to dynamic speed limit control to eliminate shock waves on freeways that is based on shock wave theory, the so-called SPECIALIST algorithm (SPEed ControllIng ALgorIthm using Shock wave Theory) [1], [2]. It was shown by simulation that the algorithm is capable of resolving shock waves, and that the improvement of the total time was that the vehicles spend on the freeway is comparable with approaches using other control techniques [3], [4].

On freeways basically two types of traffic jams can occur: jams with the head fixed at a bottleneck location and jams that have an upstream moving head and tail. Here we focus on the second type, which are often called shock waves [3] or wide moving jams [5]. We will use the term shock wave for these jams. These jams are typically short jams (say, 1-2 km) that propagate upstream, due to the incoming vehicles at their tail and the leaving vehicles at their head. They can remain existent for a long time and distance [5]. Consequently, every vehicle that enters the freeway upstream of the jammed area will have to pass through the jammed area, which increases travel times, creates potentially unsafe situations, and increases noise and air pollution by braking and accelerating vehicles. Shock waves typically have a significantly lower outflow than the capacity of the freeway, which motivates the idea that traffic flow can be improved by resolving shock waves. The difference between the free flow capacity and the queue discharge rate is around 30% [5].

In the literature two main approaches can be found to dynamic speed limit control aiming at flow improvement. The first emphasizes the homogenization effect [6]–[9], whereas the second is focused on preventing traffic breakdown or resolving existing jams by reducing the flow by means of speed limits [3], [4], [10], [11]. The basic assumption of homogenization is that speed limits can reduce the speed (and/or density) differences, by which a more stable (and safer) flow can be achieved. The flow reduction approach focuses more on preventing or resolving too high densities (including jams) by limiting the inflow to them. By resolving the jams (the bottlenecks), higher flows can be achieved in contrast to the homogenization approach as demonstrated in [3], [4], [10].

SPECIALIST is of the flow limitation type, is based on a simple principle, has a very low computational demand, and has tuning parameters that have a physical interpretation which makes the tuning parameters straightforward.

In the remainder of the paper the theory of the algorithm will be discussed in Section II. The algorithm that was developed and tested based on this theory is presented in Section III. Section IV discusses the tuning variables and the approach to the tuning procedure. The approach for the evaluation of the performance of the algorithm is discussed in Section V, followed by the description of the implementation set-up in Section VI. In Section VII the results are discussed, and the conclusions are presented in Section VIII.

II. THEORY OF SPECIALIST

The theory of resolving shock waves by dynamic speed limits is based on the shock wave theory as also applied by Lighthill and Whitham in their famous paper [12]. Before explaining the approach for shock wave resolution we explain one of the fundamental relationships in shock wave theory. Although shock wave theory goes further than what is presented here, we only present one fundamental aspect, which is necessary to understand the remainder of the paper.

A. Shock wave theory

One of the basic relationships in shock wave theory is the relationship between the time-space graph of the traffic states (as shown on the left in Fig. 1) and the density-flow graph (as shown on the right in Fig. 1). The time-space graph shows the traffic states on a road stretch (along the vertical axis) and their propagation over time (in the horizontal direction). In the figure a short traffic jam is shown...
that propagates upstream (area 2) and which is surrounded by traffic in free-flow (areas 1). The density-flow diagram shows the corresponding density and flow values for these states. Shock wave theory states that the front (boundary) between two states in the left figure has the same slope as the slope of the line that connects the two states in the right figure. Note that the slopes in both figures have the unit of \( \text{km/h} \), and that states in the right figure correspond to areas in the left figure. The orange lines (light gray in black and white) indicate the fundamental diagram (as a reference).

The importance of this relationship is that if the different traffic states on a freeway stretch are known, then their future evolution can be predicted by describing the fronts between them. This basic relationship will be used in the theory for resolving shock waves.

### B. Control approach

The approach to resolve shock waves consists of four phases and starts with a shock wave similar to the example above.

**Phase I.** Assume a shock wave (as shown in Fig. 2) is detected on the freeway. (How the shock wave is detected will be explained in Section III.) We assume that the traffic state upstream (state 6) and downstream (state 1) of the shock wave is in free flow which is generally the case in real traffic. In Fig. 2 the phases are indicated at the top of the left sub-figure. For the sake of readability of the figures we assume that state 1 and state 6 are equal, but the theory also holds for the case when they are unequal.

**Phase II.** As soon as the shock wave is detected the speed limits upstream of the shock wave are switched on. This leads to a state change in the speed-controlled area from state 6 to state 3 (in Fig. 2 approximately from 4-8 km), and to the boundary between areas 6 and 3. State 3 has the same density as state 6, as the density does not change when the speed limits are lowered on a longer stretch; no vehicles can suddenly appear or disappear. However, the flow of state 3 is lower than that of state 6 due to the combination of the same density with a lower speed.

As shown by the density-flow graph, the front between states 2 and 3 will propagate backwards with a lower speed than the front between states 1 and 2, and consequently the two will intersect and the shock wave will be resolved after some time.

At the upstream end of the speed-limited area traffic will flow into this area, with the speed equaling the speed limit and with a density that is in accordance with the speed, typically significantly higher than the density of state 3 (which was the density corresponding to free-flow). This state is called state 4. The front between states 6 and 4 will propagate upstream or downstream depending on the flow of state 4. (The density and consequently the flow is a tuning variable, which will be explained in Section IV.)

We choose the initial length of the speed-limited stretch such that the creation of state 3 exactly resolves the shock wave. This length can be determined by using point D (the intersection of the fronts 1-2 and 2-3) and the known slope of front 3-4. The length of the stretch CE depends on the density and flow associated with states 1, 2 and 6, the speed corresponding to state 3 (and consequently, the resulting flow reduction), and the physical length of the detected jam.

**Phase III.** When the shock wave (area 2) is resolved, there remains an area with the speed limits active (state 4) with a moderate density (higher than in free-flow, but lower than in a shock wave) and a moderate speed. A basic assumption in this theory is that the traffic from such an area can flow out more efficiently than a queue discharging from full congestion as in the shock wave. So, the traffic leaving area 4 will have a higher flow and a higher speed than state 4, represented by state 5. This leads to a backward propagating front between states 4 and 5, which resolves state 4.

**Phase IV.** What remains is state 5, and state 6 upstream and state 1 downstream of it. The fronts between states 1 and 5, and between states 6 and 5 propagate downstream, which means that eventually the backward propagating shock wave is converted into a forward propagating wave leading to a higher outflow of the link as shown in Fig. 2.

Obviously, not all traffic situations are suitable to construct the above control scheme. The exact requirements for such a scheme will be discussed in Section III-C when the resolvability assessment is discussed.

---

**Fig. 1.** According to the shock wave theory the propagation of the front between two traffic states in the left figure has the same slope as the line connecting the two states in the density-flow diagram in the right figure. The arrow indicates the travel direction.

**Fig. 2.** The four phases of the SPECIALIST algorithm. Phase I: The shock wave is detected. Phase II: Speed limits are turned on in areas 2, 3, and 4. The shock wave dissolves. Phase III: The speed-limited area (area 4) resolves and flows out efficiently. Phase IV: The remaining area 5 is a forward propagating high-speed high-flow wave.
III. ALGORITHM DEVELOPMENT

Based on the theory in Section II an algorithm is developed that is suited for real-world implementation. The algorithm consists of four steps: shock wave detection, control scheme generation, resolvability assessment, and control scheme application, which will be detailed in the following sections.

A. Shock wave detection

In the shock wave detection step the shock wave is detected by using thresholds \( v_{\text{max}} \) (km/h) for the speed and \( q_{\text{max}} \) (veh/h) for the flow measurements, by assuming that in segment \( i \), a shock wave is present if \( q_i \leq q_{\text{max}} \) and \( v_i \leq v_{\text{max}} \). When there are no other types of jams on the considered stretch, this identifies the location of the shock wave.

To compensate for the dependency of the head and tail location on the threshold values, and to cope with the fact that the shock wave can only be detected at discrete locations and discrete times additional offsets \( x_{\text{head-offset}} \) (km) and \( x_{\text{tail-offset}} \) (km) are defined for the estimation of the location of the head and tail. The location of the shock wave head and the tail that the algorithm uses to determine the control scheme is the sum of the detector locations corresponding to the jam and the offsets. In this way even when the jam is detected by only one detector, the jam length will larger than zero.

B. Control scheme generation

In this step all the traffic states in the control scheme are determined. The traffic states are denoted \( v_{[j]}, q_{[j]}, \rho_{[j]}, j \in \{1, \ldots, 6\} \) for the six states. Since \( q_{[j]} = \rho_{[j]} v_{[j]} \), it is sufficient to determine two of the three variables for each state. The areas directly upstream and downstream of the shock wave that are not falling below the thresholds \( q_{\text{max}} \) and \( v_{\text{max}} \) are classified as free-flow, and the measurements from these areas are determined to calculate the average states upstream and downstream of the shock wave.

The average flow \( q_{[6]} \) and \( q_{[1]} \) (veh/h/lane) and average density \( \rho_{[6]} \) and \( \rho_{[1]} \) (veh/km/lane) respectively upstream and downstream are determined by \( q_{[6]} = (1/N) \sum_{i \in I_{[6]}} q_i \), and \( \rho_{[6]} = (1/N) \sum_{i \in I_{[6]}} \left( q_i/v_i \right) \), (and similarly for state 1) where \( q_i \) (veh/h/lane) and \( v_i \) (km/h) are respectively the flow and speed measurements of segment \( i \), and \( I_{[6]} \) and \( I_{[1]} \) is the set segment indexes of the segments in free flow respectively upstream and downstream from the jam, and \( N \) the number of segments considered. This determines states 1 and 6.

The density of state 2 is determined in a different way since typically a very low speed is associated with state 2 and induction loop detectors are known to be inaccurate for low speeds. The density of state 2 is determined by using the density and flow of state 1, the flow of state 2, and solving \( \rho_{[2]} = \left( q_{[2]} - q_{[1]} \right) / \left( v_{[1]} - v_{[2]} \right) \), for \( \rho_{[2]} \), where \( v_{[1,2]} \) denotes the propagation speed of the head of the shock wave. The propagation speed of the head of the shock wave can be taken from off-line data, since it is well-known to be fairly constant (around -18 km/h). State 3 directly follows from the density of state 6 and the used speed reduction. The speed of state 4 equals the value of the speed limits. However, the density of state 4, and the density and flow of state 5 do not follow from the shock wave theory and can be considered as a design variables for which heuristic tuning rules can be given. For now, it is sufficient to assume that they have a fixed value. Tuning will be discussed in Section IV.

When all six states are determined, the control scheme can be constructed. This involves the determination of the various fronts and their intersection points by solving straightforward linear equations based on the traffic states and shock wave theory. Due to space limitations we do not present the equations here.

C. Resolvability assessment

After the construction of the control scheme the resolvability is assessed. A shock wave is as resolvable if:

1) The control scheme can be constructed given the measured and calculated traffic states 1–6. This means that the head and tail of area 2 should converge, otherwise the shock wave will not be resolved. The same applies for area 4.

2) State 5 should has higher flow and a higher density than state 1, otherwise there is no forward propagating front between these two states. Furthermore, the speed of state 5 should be less than or equal to the speed of state 1 in order to preserve the general shape of the fundamental diagram in free flow.

3) The speed in area 6 is higher than the speed limit, otherwise the speed limits do not have any effect.

4) The necessary length of the speed-limited stretch is smaller than the available upstream free-flow area.

If these conditions are satisfied, the shock wave is classified as resolvable and the control scheme is applied. Otherwise, the algorithm waits for new data.

D. Control scheme application

If the control scheme is determined and the shock wave is resolvable then the speed limit can be activated. For each gantry location it is determined at what time it falls in the area’s 2, 3, or 4 (where the speed limit should be active).

IV. TUNING

In the algorithm there are nine parameters that can be selected by the designer of the control system. For these variables heuristic tuning rules can be given, partially based on offline traffic data and partially based on the on-line (closed-loop) behavior of the algorithm. Below we discuss the tuning approach for these variables. The actual changes that were made to tune the parameters will be discussed in Section VII together with the results. Due to spacial limitation we only discuss the parameters here that were changed during the test.

One of the most important tuning parameters is the density associated with state 4. The speed of state 4 is determined by the speed limits, however the choice of the density is a design variable that influences the shape of the control scheme. If the density is chosen higher, the slope between states 4 and 6 will be less steep, which means that the tail of
the speed limits will propagate less quickly backwards. This relationship also can be understood the other way around: by letting the tail of the speed limits (the front between states 4 and 6) propagate faster backwards, the density in area 4 can be kept low.

The relevance of the density is that traffic traveling at a given speed is stable at sufficiently low densities and becomes more and more unstable with increasing densities. By selecting 4 properly (low enough), the stability of the traffic under the speed limits can be ensured.

Other tuning parameters are the speed \( v[5] \) and flow \( q[5] \) of state 5. These parameters determine the slopes 1–5, 4–5 and 5–6. Slope 4–5 determines how the speed limits are release at the head of the speed-limited area. By changing the speed of the slope the flow of state 5 can be influenced. Obviously, the resulting flow cannot take any value, due to the autonomous behavior of vehicles leaving area 4 (acceleration, lane-changing, etc.). Nevertheless, within the bounds of “physical” possibilities, the choice of \( q[5] \) will influence the outflow resulting from the control scheme. This outflow should be not too high to prevent triggering new jams at downstream bottlenecks. Furthermore, \( q[5] \) should be higher than \( q[4] \) and \( \rho[5] \) should be lower than \( \rho[4] \) in order to have a backward propagating front.

V. EVALUATION APPROACH

Both traffic performance and the operation of the algorithm in real traffic is evaluated. We discuss the evaluation questions and the approach for answering these questions. In the approach the visual assessment of the traffic scenario in the space-time plots of the speed, flow and density plays an important role. Although this method is not rigorous, the information in these plots was invaluable for the tuning and evaluation of the algorithm. Currently there are no systematic methods available that can reliably distinguish between several types of jams, such as shock waves (moving jams), jams at on-ramps, jams caused by accidents, etc.

A. Traffic performance evaluation

For the evaluation of the traffic performance the following performance indexes were considered:

- The number of resolved jams. For each speed limit intervention the targeted jam was evaluated whether is was resolved by the speed limits. Resolution can be identified by disappearance of the high density and very low flow that is associated with shock waves. An example of a successful resolution is given in Fig. 3 and of an unsuccessful resolution in Fig. 4.

- The gain in veh-hours. The gain in time that the vehicles spend in the considered section is calculated for each intervention in the following way. First, the basis of comparison is the assumption that in case of no intervention the jam would have lasted for an hour, and would have had an outflow of 1500 veh/h/lane. These assumptions were taken from other studies that were performed previously on this freeway stretch. Next, the flow for each intervention is measured at 2 km downstream of the line AD for the duration of the control scheme (from A to F). The distance of 2 km was chosen such that the vehicles leaving the jam are not accelerating anymore (during acceleration the full flow is not reached) but not too far downstream to prevent measuring flows corresponding to other jams. From the difference in flows of the assumed uncontrolled scenario and the measured flows the gain in vehicle-hours was determined.

- New jams triggered. Each intervention was judged whether new jams were created in or upstream of area 4 (which may be due wrong tuning of the density 4) and downstream (which may be due to the tuning of state 5). An example where three upstream jams were created is given in Fig 5.

- Compliance. The compliance of the traffic with the shown 60 km/h is determined by investigating the average and the standard deviation of the speed in areas 3 and 4 for each intervention.

B. Evaluation of the algorithm

The algorithm has been evaluated whether the general operation is as expected, and whether the scheme has the expected shape. Furthermore, by visual inspection of the space-time plots the type of jam was determined for which the speed limits became active. These types were classified as (1) shock wave, or (2) other type, where the other types of jams include jams at on-ramps (recognizable in the space-time plots by the fixed head at an on-ramp location), fixed jams at a bottlenecks (fixed head, but no on-ramp), jams due to accidents (non-structural location combined with very low flow), a few cases of jams caused by snowplows (forward moving bottleneck combined low flow), and jams of unclear type.

VI. EXPERIMENT SET-UP

The freeway stretch for the test is a part of the Dutch A12 freeway and has three lanes and a length of approximately 14 km. The stretch is located between the connection with
the N11 at Bodegraven up to Harmelen as shown in Fig. 6. The stretch includes a three on-ramps: at Bodegraven, at Nieuwerbrug, and at Woerden. From offline data analysis it was known that the shock waves are often created around km 48 (in the figure behind the green E25 sign).

The stretch is equipped variable message signs that displays the speed limit. The spacing of the gantries is 500 to 600m. Under each gantry there are double loop detectors (one pair for each lane), measuring the speed and flow. The speeds and flows were averaged over a period of one minute, but with a sampling time of 10s. At some locations there were additional loop detectors which were also used for the traffic measurements.

The displayed speed limit for this algorithm was 60 km/h, and for the lead-in 80 and 100 km/h were used.

VII. RESULTS

In this section we discuss the performance of the SPECIALIST algorithm and the tuning steps taken during the test. Since the tuning steps depended on the performance results they will be discussed together in Section VII-A. In Section VII-B we discuss the findings regarding the operation of the algorithm.

A. Traffic performance and tuning

The initial tuning parameters were selected based on experience with the offline data and the simulation studies, and are shown in Table I. Table II shows the performance indexes discussed in Section V. Since the algorithm also activated for other types of jams than shock waves only, the table distinguishes the results according to jam type.

From Table II it is clear that in the period of the third parameter setting the performance was significantly lower than in the other periods. This is probably due to the extremely bad weather in this period, where there was a lot of snowfall. In this period often one or two lanes were closed and snowplows were used on a regular basis.

The algorithm activated in roughly the half of the cases for shock waves, and the other half for other jam types. For the shock waves the average gain in total travel time was 35 veh-h, while for the other jams it was close to 0 veh-h.

During the test the density \( \rho_{\text{[4]} \text{veh/km/lane}} \) was lowered gradually, in order to stabilize traffic and reduce the number of new jams appearing upstream of the speed-limited area. This led to a reduction from 0.33 to 0.07 new upstream jams.

The compliance during the whole test was roughly constant and was somewhat better (lower) in area 4 than in area 3. This might be due to the higher density in area 4, which may motivate drivers better to drive slower.

B. Algorithmic results

- The algorithm activated frequently for other jams than shock waves. This is mainly due to the simple way that shock waves are detected and the desire to detect shock waves as early as possible. Early detection of shock waves entails that there is not time to track the head of the jam and see if it is backwards propagating (which is the distinguishing feature of shock waves).
- In a relatively large number of cases the shock waves were classified as not resolvable due to the resolvability condition 2 in Section III-C. The reason was that (the
measured) $v_1$ was lower than $v_5$ or that the $\rho_1$ was higher than $\rho_5$. These cases were due to the random fluctuations in the downstream state. Since for a given state 4, there are several solutions for state 5 that result in the same slope 4-5 (all the solutions are on the same line), another setting was chosen for state 5, which had a lower speed and a higher density (but with the same slope 4-5) in order to minimize the number of shock waves that are assessed as not resolvable. The effect of this change on the number of activations per day can be seen in Table II.

VIII. CONCLUSIONS

The real-world test results of the SPECIALIST algorithm are discussed in this paper. The major finding is that it is demonstrated in practice that it is possible to limit the inflow of a traffic jam by dynamic speed limits while keeping the traffic flow stable. Regarding the quantitative results, about 80% of the shock waves that were theoretically resolvable, were resolved in practice by the algorithm, and the average gain of total travel time per resolved shock wave was 35 vehours. In the majority of the cases the algorithm generated control schemes according to the (qualitative) expectation. For the few cases when the behavior of the algorithm other than expected, solutions have been suggested.

ACKNOWLEDGMENTS

Supported by the DYNAMAX project of the Dutch Ministry of Transport, Public Works and Water Management.

REFERENCES


TABLE I

<table>
<thead>
<tr>
<th>parameter setting</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td># days</td>
<td>49</td>
<td>42</td>
<td>35</td>
<td>23</td>
<td>149</td>
</tr>
<tr>
<td># activ. total</td>
<td>84</td>
<td>60</td>
<td>31</td>
<td>67</td>
<td>242</td>
</tr>
<tr>
<td>av. activ. per day</td>
<td>1.7</td>
<td>1.4</td>
<td>0.9</td>
<td>2.9</td>
<td>1.6</td>
</tr>
<tr>
<td>resolved %</td>
<td>68%</td>
<td>60%</td>
<td>52%</td>
<td>58%</td>
<td>59%</td>
</tr>
<tr>
<td>av. new jams upstr. per activ.</td>
<td>0.13</td>
<td>0.18</td>
<td>0.13</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>av. new jams downstr. per activ.</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>av. gain veh-h per activ.</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>26</td>
<td>18</td>
</tr>
<tr>
<td>shock waves</td>
<td>48</td>
<td>24</td>
<td>11</td>
<td>32</td>
<td>115</td>
</tr>
<tr>
<td>resolved</td>
<td>83%</td>
<td>83%</td>
<td>55%</td>
<td>72%</td>
<td>71%</td>
</tr>
<tr>
<td>av. gain veh-h per activ.</td>
<td>30</td>
<td>41</td>
<td>33</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>other jams</td>
<td>36</td>
<td>36</td>
<td>20</td>
<td>33</td>
<td>127</td>
</tr>
<tr>
<td>resolved</td>
<td>47%</td>
<td>44%</td>
<td>20%</td>
<td>46%</td>
<td>41%</td>
</tr>
<tr>
<td>av. gain veh-h per activ.</td>
<td>-2</td>
<td>-1</td>
<td>-8</td>
<td>14</td>
<td>2</td>
</tr>
</tbody>
</table>