WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
*Wirel. Commun. Mob. Comput.* 2013; **13**:266–276
Published online 26 January 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2200

SPECIAL ISSUE PAPER

# A new traffic congestion prediction model for advanced traveler information and management systems

Gustavo Marfia, Marco Roccetti* and Alessandro Amoroso

Mura Anteo Zamboni 7, Computer Science Department, University of Bologna, 40127 Bologna, Italy

## ABSTRACT

The increase of wasted time and pollution due to vehicular traffic has paved the way to many different countermeasures, ranging from the enforcement of congestion tolls to the commercialization of vehicles powered by low-emission hybrid engines. Advanced traveler information systems (ATISs), which are capable of supplying updated traffic information to all those citizens that are driving through city roads, represent a prominent approach to combat vehicular congestion. In brief, ATISs are concerned with collecting, processing, and disseminating traffic information, providing data that can be profitably exploited by an on-board navigation system to compute the most convenient route to a given destination. Indeed, their role becomes progressively more relevant as their accuracy and reliability increases, thus encouraging more and more people to utilize them while driving. With this in mind, we devised a new congestion detection model that accurately estimates and forecasts the short-term congestion state of a road without requiring any prior knowledge regarding any of its parameters. Such model can be easily integrated within an ATIS and usefully applied to any given road. The efficacy of our model is here proved through the results of several experiments, which witness the validity of our approach. Copyright © 2012 John Wiley & Sons, Ltd.

**KEYWORDS**

vehicular congestion; advanced traveler information systems; vehicular mobility models; intelligent transportation systems

***Correspondence**

Marco Roccetti, Mura Anteo Zamboni 7, Computer Science Department, University of Bologna, 40127 Bologna, Italy.
E-mail: roccetti@cs.unibo.it

## 1. INTRODUCTION

Vehicular congestion in urban areas has steadily grown during the years to become one of the primary problems tackled by city administrators. In the USA, for example, statistics collected from its 15 most populated cities show that the amount of time wasted by daily commuters has more than doubled in the past 30 years, reaching the value of more than 50 h, on average, per year [1]. In developing countries, as China for example, this trend is even more dramatic, as the rate at which motor vehicle ownership increases is orders of magnitude higher than the rate at which new roads are being constructed [2]. The constant increase of vehicular congestion in many cities, throughout the world, has led to the devise of a plethora of different approaches that aim at stopping and, possibly, inverting such trend. One of such approaches, termed advanced traveler information systems (ATISs), is built upon the joint use of advanced information and communication technologies and global positioning systems (GPSs). In brief, wireless communication infrastructures and navigation technologies are exploited to (i) collect positions from vehicles, (ii) infer relevant information concerning vehicular congestion, and (iii) provide congestion information to each single driver.

All these are possible thanks to a widespread deployment of advanced personal navigation devices, which are equipped with wireless access interfaces (e.g., Universal Mobile Telecommunications System, WiFi, WiMax), a GPS receiver, and a digital map. Additionally, the implementation of such type of systems requires the use of a central entity that receives, stores, and processes vehicle positions and, in turn, sends updated congestion estimates and forecasts to all drivers, as represented in Figure 1. Once drivers discover that one of the road sections they planned to traverse is severely jammed, they can decide to change their route before remaining stuck in traffic.

Beneficial applications of ATISs go well beyond the vehicular congestion domain, including also, but not being limited to, important contributions that they can bring to


Copyright © 2012 John Wiley & Sons, Ltd.

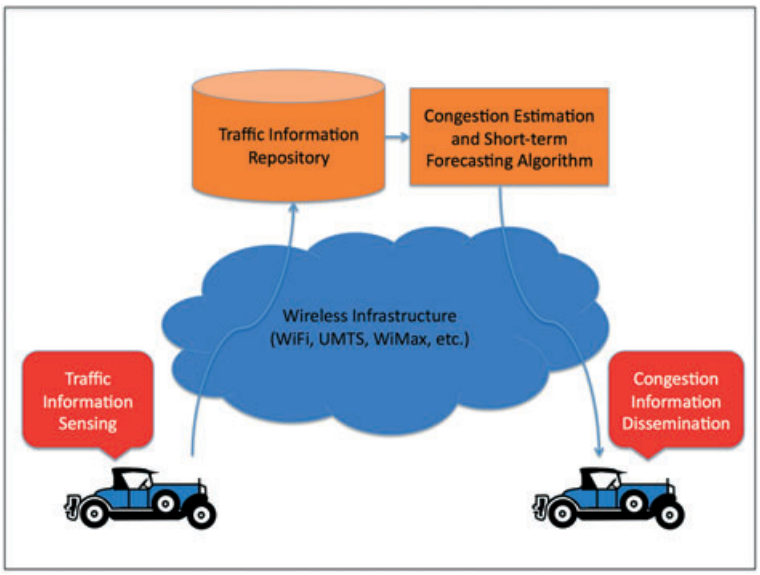


**Figure 1.** Scenario.

environmental, drivers, and pedestrian safety, for example [3–11]. In fact, a reliable knowledge of both the real-time and future traffic situation could, for example, aid city traffic managers in taking operative decisions concerning the diversion of traffic flows, when congestion levels in given areas exceed certain pre-defined thresholds. An efficient ATIS infrastructure could also offer a valuable tool to detect or even foresee accident occurrences, for example, alerting when traffic very rapidly builds up, thus indicating the happening of a non-recurrent event (e.g., an accident), or when traffic flows stop following given known patterns (e.g., vehicles entering a one-way street in the wrong direction), thus predicting a dangerous event might take place.

Clearly, the successful deployment of the aforementioned systems heavily depends on the reliability with which congestion estimates and forecasts are performed. Wrong estimates and forecasts, in fact, could divert traffic flows towards congested areas, thus prolonging the time required to reach a given destination.

Although the vehicular congestion phenomenon has been thoroughly studied for highway scenarios, where identifying traffic conditions is typically simple because a limited number of factors (e.g., incoming traffic flows and obstacles) contribute to a given traffic situation that may be found at any given time (possible causes of congestion are marked with flames in the leftmost part of Figure 2), the same cannot usually be said for urban scenarios [12,13]. In fact, in urban areas, several can be the underlying causes of traffic jams (rightmost part of Figure 2), which can hence more hastily appear and disappear and consequently be more difficult to detect and predict. Hence, the initial step that must be undertaken in engineering an ATIS that may prove to be effective in urban areas is the devise of an operative definition that clearly indicates when vehicular congestion appears on any given road segment.

Indeed, many different definitions have been presented in literature, but, to the best of our knowledge, none provides results that are not prone to interpretation while being independent from specific road parameters. For this reason, we propose a new traffic congestion definition, which draws its inspiration from a relevant line of research that has studied how the capacity and the bandwidth occupied by an Internet connection could be efficiently determined by observing the end-to-end delay of Internet Protocol packets [14,15].

Our definition starts from the observation that any two cars that traverse the same street will probably experience the same congested traffic conditions, assuming that they do traverse it not too far away in time one from the other. This is due to the inertia of vehicular queues, which causes a road to remain congested for a given period equal to, say $S$, even if, at a certain point in time, the ingress flow to that road rapidly drops. Hence, such observation allows us to provide the definition of congestion, which is as follows: congestion is a road's state, which *lasts for at least S units of time*, during which travel times exceed the time $T^*$ normally incurred under light or free-flow travel conditions.

Our aim, in this paper, is to introduce a new congestion detection and short-term forecasting algorithm, which, on the basis of the congestion definition mentioned before, can be easily put to good use within ATISs.

The strength of our algorithm is that it can provide its results without requiring any prior knowledge regarding a road. In fact, our algorithm is able to return short-term congestion forecasts by simply processing the information collected by the personal navigation device units of vehicles. In order to verify the validity of our approach, we devised a wide set of experiments that led us to drive for over 450 miles on urban sections of Los Angeles (CA) and Pisa (Italy) roads.

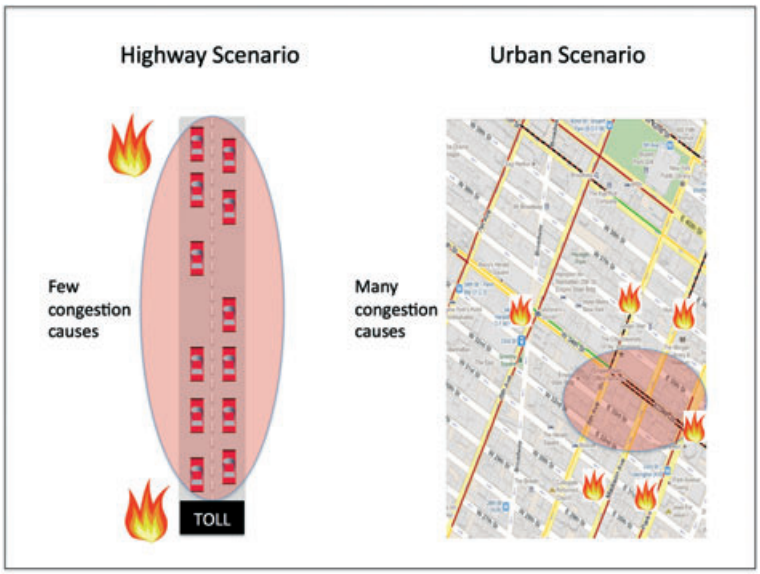


**Figure 2.** Highway versus urban congestion.

This paper is organized as follows. In Section 2, we survey the schemes that fall closest to ours in recognizing congested from non-congested roads. In Section 3, we briefly describe the model underlying our algorithm and verify its robustness in Section 4, whereas we demonstrate how it can be beneficially employed within an ATIS in Section 5. The results of our experimental assessment are provided in Section 6. Section 7, finally, concludes our paper.

## 2. RELATED WORK

Although a wealth of work in the area of congestion detection and forecasting algorithms exist, for the sake of brevity, we here only survey three different approaches that most easily can be integrated with an ATIS [13–22].

The first one, known as Surface Street Traffic Estimation, was introduced to recognize congestion situations on streets that are controlled by traffic lights [20]. Succinctly, according to this approach, vehicles are considered to undergo a congestion event if they step in one of the two situations that follow that decelerate their flow. In fact, vehicles could experience delays either by moving in a *stop and go* pattern or by simply standing in queue for a full red light cycle or more. The shortcoming of this approach is that it lacks any short-term forecasting power and, to the best of our knowledge, it has only been defined and tested on streets that end at signalized intersections, thus leaving out of its scope and applicability a wide class of urban streets.

The second approach we tested here has been drawn from the Highway Capacity Manual (HCM) delay formula for signalized intersections [21]. This formula computes the average traversal time $T_{\mathrm{HCM}}$ a vehicle experiences on a given street, which varies with the street length, speed limit and capacity (as a function of its number of lanes and its traffic light phases length), and the average amount of cars that enters during a given time span. $T_{\mathrm{HCM}}$ returns useful information to recognize when a given section of a road is congested or not, when computed, setting the ingress traffic volume equal to the road's capacity. Although very interesting, such approach is limited by the fact that it depends upon prior information, such as given road parameters, which hence jeopardize its ability to adapt to new situations.

To conclude with the third proposal, we refer to the service provided by Google Traffic, as a practical service of common use providing drivers with traffic information before they begin their journey (through the Web) and while driving, through navigation units that are installed on their cellphones [22]. Although such system is gaining an increasing role in the management of urban platoons of vehicles, little information has been so far published on how it really works and on how it actually detects a congested state on a given road. Not only that, it generally lacks any information concerning the trend of the traffic situation on a given road in the short term, thus omitting one of the pieces of the puzzle that most may influence the routing decisions taken by drivers. All these said, from the point of view of a user, a color is assigned to each road, depending on the average speed that has been measured on it and on its speed limit. For highways, for example, green represents an average speed that exceeds 80 km/h, yellow an average speed that lies between 40 and 80 km/h, and red less than 40 km/h.

For the sake of completeness, we listed the main characteristics (pitfalls) of the schemes discussed in this section in Table I.

Differently from any of the proposals that have been described here, we devised an algorithm that can

**Table I.** Congestion detection tools: characteristics.

| | Forecasting | Detection | Use of road parameters |
|---|---|---|---|
| Surface Street Traffic Estimation | No | Yes | No |
| Highway Capacity Manual | No | Yes | Yes |
| Google Traffic | Yes (based on historical information) | Yes | No |

contemporarily detect and forecast, in the short term, congestion on any given road, not needing any prior information regarding any of its parameters.

## 3. DETECTING AND FORECASTING TRAFFIC CONGESTION: A NEW APPROACH

With the traffic congestion definition anticipated in Section 1, we now provide a method to compute the congestion threshold $T^*$, as well as the time $S$ for which congested or non-congested conditions last on any given road $R$. The scheme that computes these values is as follows.

We consider a road $R$ as congested if it is possible to find a value of $T^*$ for which, when a vehicle spends at least time $T^*$ on it, the majority of subsequent cars (e.g., 80%) that later enter $R$ (say within a time span $S$) still spend at least $T^*$ units of time to leave it. If, instead, only a low percentage of the cars that later entered that road experienced a traversal time above $T^*$ units of time (e.g., 20%), this entails that $R$ is leaving a state of congestion.

In a similar manner, $R$ is non-congested when, if a car takes less than $T^*$ units of time to leave it, the majority of vehicles that later enter it (e.g., again 80%) still require less than that time. Obviously, in the case that only a low ratio of vehicles that later entered that road experienced a traversal time below $T^*$ (e.g., 20%), this would mean that $R$ is transitioning into a congested state. The 80% value is drawn from the literature, but it may depend on the specific road under consideration [23].

It is now our duty to translate the aforementioned considerations into a more formal modeling setting, which will then allow us to derive our congestion detection and short-term forecasting algorithm.

To this aim, we now introduce the four sets of vehicles: (i) $HC(T_1^*)$, the number of pairs of vehicles that suffer from high congestion, (ii) $N1(T_1^*)$, the number of pairs of vehicles that are leaving a congested situation, (iii) $NC(T_2^*)$ the number of pairs of vehicles that do not suffer from congestion, and, finally, (iv) $N2(T_2^*)$ the number of pairs of vehicles that are entering a congestion state. Specifically, we have the following definitions.

**Definition A (High Congestion Set)**. Consider a group $P$ of vehicles entering a street $R$, with the first vehicle of the group entering $R$ at a given time and the last one entering $R$ at $S$ units of time later than the first one. $HC(T_1^*)$ is defined as the set of all the pairs of vehicles, $(i, j)$, in $P$ for which both their traversal times, say $T_i^*$ and $T_j^*$, exceed the congestion threshold $T_1^*$. We also define as $N1(T_1^*)$ the set of all the pairs of vehicles, say $(h, k)$, in $P$ for which the traversal time $T_h^*$ of only the first vehicle $h$ exceeds $T_1^*$.

**Definition B (Low Congestion Set)**. Take the same group of cars $P$ entering $R$ within a time span of the same length as before. $NC(T_2^*)$ is defined to be the set of all the pairs of cars, say, $(i, j)$, in $P$ for which both their traversal times, say, $T_i^*$ and $T_j^*$, are below the congestion threshold $T_2^*$. Consequently, $N2(T_2^*)$ is the set of all the pairs of cars, say, $(h, k)$, in $P$ for which the traversal time $T_h^*$ of only the first vehicle $h$ is below $T_2^*$.

We need now to measure the size of the four aforementioned sets. To this aim, we provide an indicator function for each of the four sets of vehicles of interest, as specified in Tables II, III, IV, and V. The meaning of these indicator functions is clear, as each of these functions may be exploited to check if a given pair of vehicles $i$ and $j$ is in one of the four given congestion states (high congestion, no congestion, leaving, and entering). For example, if two given vehicles $i$ and $j$ both experience a delay exceeding the congestion threshold $T_1^*$, the $I_{HC}(i, j)$ indicator function will be equal to 1.

**Table II.** Congestion state indicator functions: high congestion.

| High congestion state |
|---|
| $I_{HC(T_1^*)} : (P \times P) \rightarrow \{0, 1\}$, where |
| $I_{HC(T_1^*)}((i, j)) = 1$, if $(i, j) \in HC(T_1^*)$ |
| $I_{HC(T_1^*)}((i, j)) = 0$, if $(i, j) \notin HC(T_1^*)$ |

**Table III.** Congestion state indicator functions: leaving congestion.

| Leaving congestion state |
|---|
| $I_{N1(T_1^*)} : (P \times P) \rightarrow \{0, 1\}$, where |
| $I_{N1(T_1^*)}((i, j)) = 1$, if $(i, j) \in N1(T_1^*)$ |
| $I_{N1(T_1^*)}((i, j)) = 0$, if $(i, j) \notin N1(T_1^*)$ |

**Table IV.** Congestion state indicator functions: no congestion.

| No congestion state |
|---|
| $I_{NC(T_2^*)} : (P \times P) \rightarrow \{0, 1\}$, where |
| $I_{NC(T_2^*)}((i, j)) = 1$, if $(i, j) \in NC(T_1^*)$ |
| $I_{NC(T_2^*)}((i, j)) = 0$, if $(i, j) \notin NC(T_1^*)$ |

**Table V.** Congestion state indicator functions: entering congestion.

| Entering congestion state |
|---|
| $I_{N2(T_2^*)} : (P \times P) \rightarrow \{0, 1\}$, where |
| $I_{N2(T_2^*)}((i, j)) = 1$, if $(i, j) \in N2(T_2^*)$ |
| $I_{N2(T_2^*)}((i, j)) = 0$, if $(i, j) \notin N2(T_2^*)$ |

**Table VI.** Congestion versus no congestion.

| |
|---|
| Road $R$ is congested if |
| $\frac{\sum_{(i,j)\in P\times P} I_{HC(T_1^*)}(i,j)}{\sum_{(i,j)\in P\times P} I_{HC(T_1^*)}(i,j)+\sum_{(i,j)\in P\times P} I_{N1(T_1^*)}(i,j)} \times 100\% \geq 80\%.$ |
| Road $R$ is non-congested if |
| $\frac{\sum_{(i,j)\in P\times P} I_{NC(T_2^*)}(i,j)}{\sum_{(i,j)\in P\times P} I_{NC(T_2^*)}(i,j)+\sum_{(i,j)\in P\times P} I_{N2(T_2^*)}(i,j)} \times 100\% \geq 80\%.$ |

It is now possible to count the ratio between the number of pairs of vehicles that traversed a given road, all taking more than $T_1^*$, and the total number of pairs of vehicles that traversed that road with the first vehicle experiencing a traversal time higher than $T_1^*$, as this value, when high, indicates a state of stable congestion for a given road $R$ (left term of the top equation of Table VI).

Similarly, the left term of the equation at the bottom of Table VI gives the ratio between the amount of pairs of vehicles that traversed a given road, taking less than $T_2^*$ units of time, and the total number of pairs where the first vehicle took a time less than $T_2^*$, as this value, when high, indicates a state of no congestion at all.

With the tools provided so far, we can verify whether a given road $R$ is congested (or not), for a given period $S$, simply by finding the values of $T_1^*$ and $T_2^*$ for which the formulas of Table VI are satisfied.

To this aim, we devised an algorithm that is based on the idea of finding the values of the thresholds $T_1^*$ and $T_2^*$, which maximize the size of the $HC(T_1^*)$ and $NC(T_2^*)$ sets (congested and non-congested states, respectively) and minimize the size of the $N1(T_1^*)$ and $N2(T_2^*)$ sets (noisy states) while avoiding the situations where this clustering activity provides spurious results due to mathematical inconsistencies. This can be performed solving the optimization problem described in Table VII.

## 4. ASSESSING THE METHODOLOGY OF THE STUDY

To assess the validity of our methodology, we have input to the equation of Table VII 100 pairs of vehicles at a time, observing the congestion thresholds it returned as the balance between congested and non-congested pairs of vehicles changed.

In essence, we computed $(T_1^*, T_2^*)$ for a given road when the numbers of pairs of vehicles that fell into its congested and non-congested sets were as follows ($NC$ identifies the set of pairs of traversal times both below the threshold of 50 s and $HC$ the set of pairs of traversal times both above 50 s): $\{|NC| = 10, |HC| = 90\}, \{|NC| = 20, |HC| = 80\}, \{|NC| = 30, |HC| = 70\}, \{|NC| = 40, |HC| = 60\}, \{|NC| = 50, |HC| = 50\}, \{|NC| = 60, |HC| = 40\}, \{|NC| = 70, |HC| = 30\}, \{|NC| = 80, |HC| = 20\}, \{|NC| = 90, |HC| = 10\}$. For each

**Table VII.** Congestion detection formula.

$$
\begin{aligned}
&(T_1^*, T_2^*) = (T_1, T_2) \text{ s.t.} \\
&\quad \{\max_{T_1^*, T_2^*} \textstyle\sum_{(i,j)\in P\times P}(I_{HC(T_1^*)}(i, j)+ \\
&\quad +I_{NC(T_2^*)}(i, j)+ \\
&\quad -I_{N1(T_1^*)}(i, j)+ \\
&\quad -I_{N2(T_2^*)}(i, j))+ \\
&\quad -|\textstyle\sum_{(h,k)\in P\times P} I_{HC(T_1^*)}(h, k) - \sum_{(y,z)\in P\times P} I_{NC(T_2^*)}(y, z)|\}.
\end{aligned}
$$

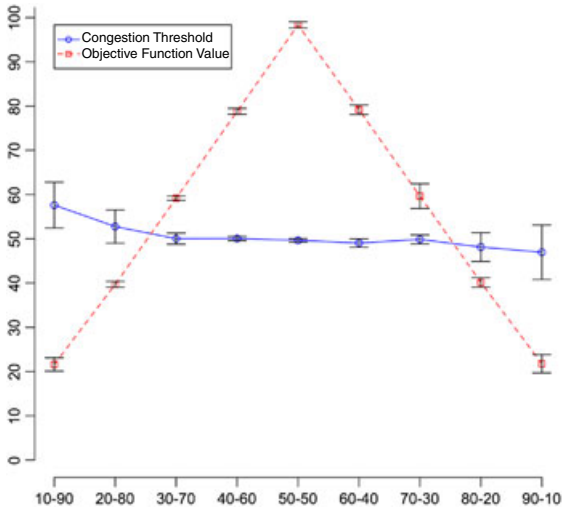


**Figure 3.** Congestion threshold and objective function, as a function of the number of pairs of vehicles that falls into the congested and uncongested sets, respectively.

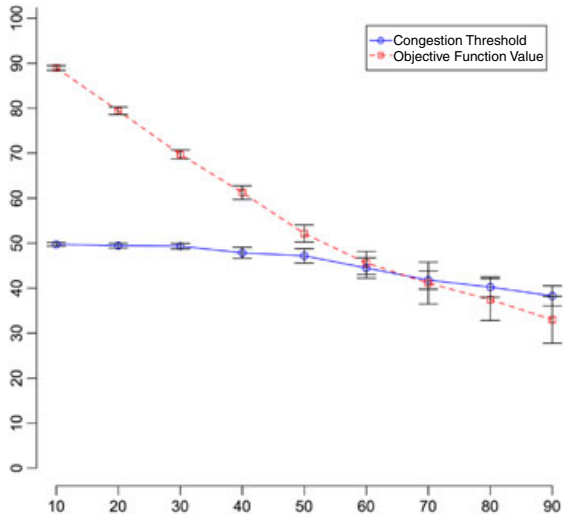


**Figure 4.** Congestion threshold and objective function, as a function of the number of pairs of vehicles that falls into the entering congestion set.

given combination, we performed 100 experiments, taking traversal times (above or below the 50-s threshold) randomly from a uniform distribution. In Figure 3, we plotted the average value of $T^*$ (which represents $T_1^*$, as $T_1^*$ and $T_2^*$ were typically very close) and that of the objective function in Table VII as a function of the previously defined combinations. Observing $T^*$, we verify that it was equal to 50 when the average value of the objective function reached its maximum and there was an equal balance between the number of pairs of vehicles that fell in the two sets (50–50). Remarkably, even when the number of pairs of vehicles that fell into the two sets became rather unbalanced, 90 congested versus 10 non-congested and vice versa, the equation of Table VII found a $T^*$ value that was near the correct one.

We also developed two further experiments where, keeping fixed the number of pairs of traversal times corresponding to congested and non-congested vehicles (50–50), in the first, we gradually increased the number of pairs of traversal times corresponding to vehicles that fell in the $N2$ set, whereas in the second, we gradually increased the number of pairs of vehicles that fell in both of the noisy sets ($N1$ and $N2$). In the first experiment, we expected that our algorithm would compute a threshold of $T^* = 50$ s until the point that the amount of introduced noise became too large. This was what occurred, as reported in Figure 4, where, again, we plotted both the average $T^*$ values (s) along with the average of the objective function. In the second experiment, instead, we found that, although the value of the objective function steeply decreased as the number of noisy samples increased (Figure 5), the balance between the noisy samples prevented $T^*$ from departing from the value of 50 s.

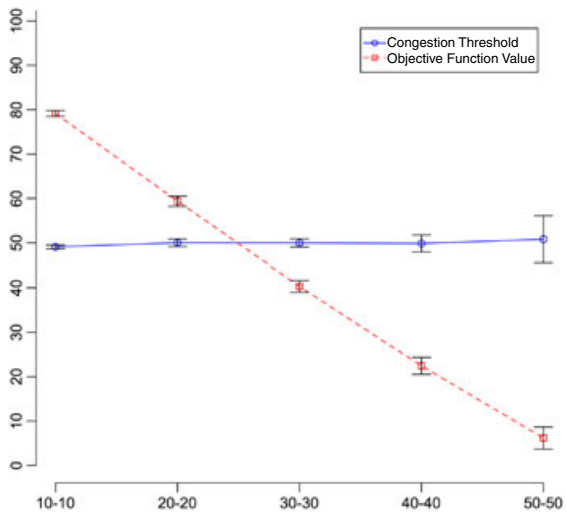


**Figure 5.** Congestion threshold and objective function, as a function of the number of pairs of vehicles that falls into the entering congestion and leaving congestion sets, respectively.

## 5. DEPLOYMENT WITHIN AN ADVANCED TRAVELER INFORMATION SYSTEM

We can put the aforementioned model to good use within an ATIS implementation adopting the following procedure.

Initially, a central entity collects traversal time data concerning a given road $R$ as returned by a set of vehicles that act as traffic probes. On receiving traversal time samples from vehicles, this entity keeps adding them to an internal data structure (line 2, Table VIII) until a sufficient number of samples have been collected and that road has been observed for a sufficient time, for example, for half

**Table VIII.** ATIS algorithm.

**Input:** Traversal time $T$ of a vehicle that traverses a given road $R$.
**Output:** Road $R$ congestion information.

```
1.if collectingData == true then
2.          collectedTraversalTimes.Add(T);
3.          if R.ObservationTime > 12 h and
4.              collectedTraversalTimes.Length > 100 then
5.                  collectingData = false;
6.                  (S, T1*, T2*) ← CTDF(collectedTraversalTimes);
7.          end
8.end
9.else
10.         if T > T*
11.                 alertCongestion(R);
12.         end
13.end
```

a day (lines 3 and 4). At the end of this initial process, the entity stops collecting information concerning the street (line 5) and builds its picture of the congestion states characterizing that given street, using the `CTDF()` function (line 6), `CTDF()` being the implementation of the formula of Table VII. If any time later, be it 1 h or 1 month, a vehicle traverses that given road exceeding the computed congestion threshold $T_1^*$, the entity can exploit this information to, for example, send a message, alerting all those vehicles that are approaching its area (line 11).

Now, it is the turn to describe how the `CTDF()` function works (Table IX). As already said, it implements the mechanism described in Table VII. Namely, it searches for the values $T_1^*$ and $T_2^*$ that contemporarily maximize the size of the high congestion set $HC(T_1^*)$ and the size of the no congestion set $NC(T_2^*)$ while minimizing the size of the remaining two sets $N1(T_1^*)$ and $N2(T_2^*)$ (lines 2 and 5). After the size of the two sets $HC(T_1^*)$ and $NC(T_2^*)$ is computed, a check is performed to evaluate whether they satisfy the conditions given in Table VI (i.e., the 80% check). If so, the function ends successfully, returning the values of $S$, $T_1^*$, and $T_2^*$. Unfortunately, the checks could fail because a too-large duration $S$ for the state of congestion of interest was chosen. This would mean that the following holds for many pairs of subsequent cars that traverse a given road: the congested (or non-congested) state a first vehicle incurs in does not last in time, as a second vehicle does not find the same state any longer. However, this problem could solely concern the value of the period $S$ that we have chosen. In fact, a smaller value of $S$ could exist, in principle, for which both of the subsequent cars that compose a vehicle pair incur in the same state of congestion. The idea is hence that of looking for such value by gradually reducing the value of $S$ until a situation is captured where both the subsequent vehicles of the pair experience a similar state of congestion (or no congestion). This motivates the iterative structure of the `CTDF()` function. As a final note, it is important to consider that our experiments show that the difference between $T_1^*$ and $T_2^*$ is always confined within a 3% value difference. This is reasonable and

**Table IX.** Congestion threshold detection function.

**Input:** A list of traversal times.
**Output:** $S$, $T^*$.

```
1.S ← Δmax minutes;
2.(T1*, T2*) ← (T1, T2) s.t. Max(T1, T2);
3.while ¬Check1(T1*) ∧ ¬Check2(T2*) ∧ S > Δmin do
4.          S ← S − δ minutes;
5.          (T1*, T2*) ← (T1, T2) s.t. Max(T1, T2);
6.end
7. if ¬Check1(T1*) ∧ ¬Check2(T2*) then
8.          return null;
9. else
10.         return(S, T1*, T2*)
11.end
```

largely expected and justifies the fact that, from now on, we will only use a unique congestion threshold value $T^*$, obtained as $T^* = T_1^* \approx T_2^*$.

## 6. TEST BED RESULTS

We carried out a set of nine different experiments in 2008 and 2009 with a fleet of cars driving through traffic to verify the effectiveness of our mobility congestion detection and forecasting algorithm. Eight of these experiments were run in Los Angeles, CA, and one in Pisa, Italy. All the main information concerning these roads are listed in Table X (name, section, length, free flow traversal time, full and green traffic light cycles, and number of traversals performed) [23,24]. Each vehicle carried an onboard system consisting of a laptop, a GPS receiver, and an EVDO interface. Upon each traversal of a given road section $R$, a car sent its traversal time to an ATIS, which, in turn, computed an estimate of $T^*$ when a sufficient amount of data was available. Our results are briefly described in the following subsection.

**Table X.** Experiment information: location, road section, road length, free flow traversal time, traffic light cycle time and green time, and number traversals performed.

| | Road | Section, direction | Length (m) | Free flow (s) | Cycle (s) | Green (s) | Lap # |
|---|---|---|---|---|---|---|---|
| 1 | Via B. Croce | Guerrazzi–Queirolo, left | 380 | 34 | 85 | 55 | 111 |
| 2 | S. Monica Blvd. | Veteran–Sepulveda, left | 380 | 61 | 120 | 15 | 134 |
| 3 | S. Monica Blvd. | Wilshire–Roxbury, straight | 280 | 17 | 90 | 54 | 77 |
| 4 | S. Monica Blvd. | Wilshire–Bedford, right | 390 | 30 | 90 | 54 | 77 |
| 5 | Lincoln Blvd. | Fiji–Venice, back | 2300 | 205 | 120 | 60 | 30 |
| 6 | Wilshire Blvd. | Midvale–Westwood, right | 130 | 7 | 150 | 80 | 71 |
| 7 | S. Monica Blvd. | Roxbury–Bedford, right | 100 | 7 | 90 | 54 | 77 |
| 8 | Wilshire Blvd. | Veteran–Westwood, right | 340 | 33 | 150 | 80 | 71 |
| 9 | S. Monica Blvd. | Westwood–Sepulveda, right | 680 | 75 | 120 | 50 | 67 |

**Table XI.** Results: congestion threshold $T^*$, $S$, $N$, and $H$.

| | $T^*$ (s) | $S$ (s) | $N$ (%) | $H$ (%) |
|---|---|---|---|---|
| 1 | 93 | 362 | 92 | 84 |
| 2 | 175 | 608 | 80 | 87 |
| 3 | 62 | 987 | 94 | 99 |
| 4 | 82 | 987 | 92 | 99 |
| 5 | 354 | 900 | 100 | 97 |
| 6 | 36 | 454 | 39 | 98 |
| 7 | 42 | 987 | 46 | 83 |
| 8 | 74 | 454 | 37 | 100 |
| 9 | 121 | 493 | 90 | 54 |

## 6.1. Results

Our results are listed in Table XI where we reported, for each street, its congestion threshold $T^*$, the duration of its congestion $S$, and the values of $N$ and $H$ (namely, the percentage of how many pairs of vehicles experienced stable congested and non-congested states, respectively).

In addition, in Figure 6, for each road number, we compare $T^*$ with two different values. The first, $\hat{T}$, represents the traversal time a car would experience when traversing that road while waiting for no more than a full red light time. The second, $T_{\text{HCM}}$, represents the average traversal times provided by the HCM method, as reported in Section 2.

Contrasting first $T^*$ to $\hat{T}$, the following observations can be made. Streets 1 through 5 all experienced alternated situations of congestion and non-congestion. The values of $N$ and $H$ in Table XI confirm this conclusion, as they both exceed the 80% threshold. Moreover, for each of these streets, the value of $T^*$ is larger than the value of $\hat{T}$ (as seen in Figure 6), which means that our algorithm correctly found congestion threshold values above which cars really experienced congestion. Streets 6 through 8, instead, were deliberately chosen because they are empirically known as almost never congested. Our results corroborate this knowledge in two different ways. First, for each of these streets, $\hat{T} > T^*$ (as seen in Figure 6), thus proving cars almost always enjoy a smooth drive because of the

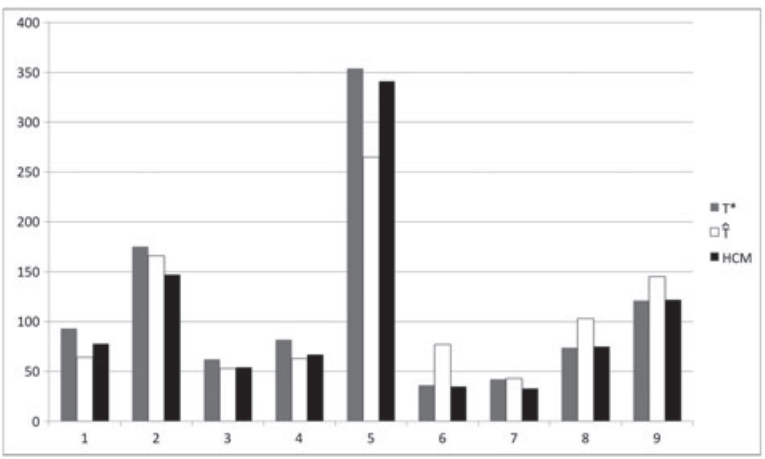


**Figure 6.** Comparison between $T^*$, $\hat{T}$, and $T_{\text{HCM}}$.

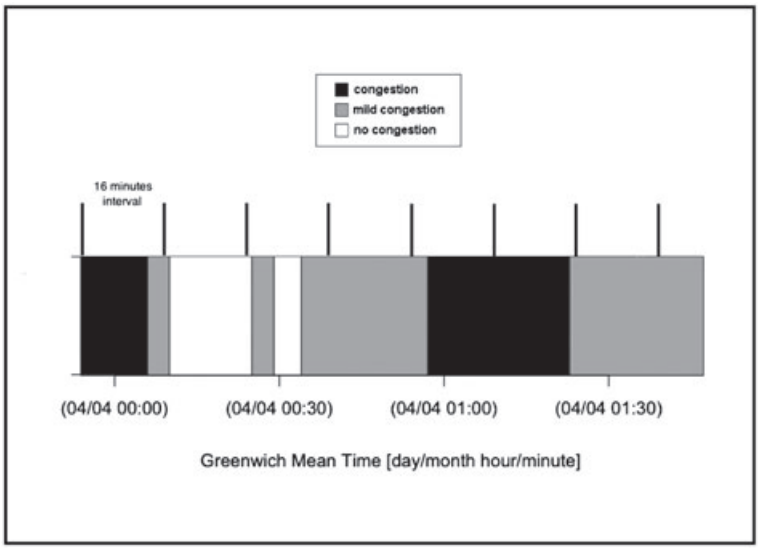


**Figure 7.** Google Traffic results for the road section of experiment 3.

existence of a green wave. Second, the very small values of $N$ (Table XI) confirm that no stable congestion was visible over those streets. Finally, street 9 requires a different discussion. A high value of $N$ and a small value of $H$ seem to reveal a stable high congestion state (Table XI). Despite this fact, $\hat{T}$ is greater than $T^*$ (as seen in Figure 6). This paradox can be explained observing that as the traffic light at the junction with Sepulveda Blvd. permits to turn right on red, only very rarely our cars stood waiting for a full red light time. To prove this with the correct explanation, we performed a few more laps with cars going straight at that intersection. As expected, during such laps the value of $T^*$ always exceeded that of $\hat{T}$.

A direct comparison of $T^*$ and $T_{\mathrm{HCM}}$ shows that the two methodologies, interestingly, return converging results on most roads. In fact, roads 5, 6, 8, and 9 give almost perfect matches (as seen in Figure 6). This fact can be easily explained, observing that those are the roads where the most accurate information regarding relevant road parameters, as the peak road capacities, provided by the Los Angeles Department of Transportation, was available [25]. When, instead, such type of information was not available and consequently default values taken from the HCM were used, we observe that the two thresholds can substantially differ (roads 1, 2, and 4).

As a final and additional experiment, Figure 7 reports the state of road 3, as provided by Google Traffic, during our experiments on the field. In particular, the three different colors (white, gray, and black) corresponded to three different traffic states (non-congested, mildly congested, and congested) that Google Traffic reported at that time. Note also the fact that the average speed below which that road was considered as congested by Google Traffic was 22 km/h.

Interestingly, comparing the results provided by Google Traffic with ours, we observed a rather good matching, both in terms of congestion threshold speeds (16 km/h) and in terms of the expected duration of congestion events. In fact, we observe in Figure 7 that periods of congestion (or no congestion) alternate, lasting approximately 15–18 min each, whereas our results revealed that congested and non-congested states last for at least 16 min ($S$ = 987 s), as depicted in Figure 7.

In conclusion, our algorithm was able to meet our expectations during our experiments, detecting when congestion occurred and estimating its minimum persistence in time. As such, and thanks to its simplicity, we believe it is the ideal candidate to be integrated in modern ATISs.

## 7. CONCLUSION

We here presented an intuitive general-purpose traffic congestion detection and short-term forecasting scheme, which has been validated on a real test bed, driving for over 450 miles along urban streets. The central part of our contribution lies in the proposal of a novel definition of congestion, where a street is defined as congested only in those cases where it has a high chance of remaining in that state for a given amount of time in the near future. This has allowed us to turn our definition into an operative algorithm that can be easily integrated into modern ATISs while resulting effective in providing relevant results of practical interest.

## ACKNOWLEDGEMENT

We are indebted towards the FIRB Damasco project for providing a financial support.

## REFERENCES


1. Texas Transportation Institute, 2010 Urban Mobility Information, [Accessed: February 17, 2011]. [Online]. Available: http://mobility.tamu.edu/ums/congestion_data/national_congestion_tables.stm..
2. Ng WS, Schipper L, Chen Y. China motorization trends, new directions for crowded cities. *The Journal of Transport and Land Use* 2010; **3**(3): 5–25.
3. Rybicki J, Scheuermann B, Kiess W, Lochert C, Fallahi P, Mauve M. Challenge: peers on wheels—a road to new traffic information systems, In *Proceedings of the Thirteenth annual ACM International Conference on Mobile Computing and Networking*, Montreal, Quebec Canada, 2007.
4. Sanwal K, Walrand J. Vehicles as probes. *Tech. Rep. UCB-ITS-PWP-95-11*, University of California, Berkeley, CA, 1995.
5. Garmin. PND-based mobile resource management solutions, [Accessed: May 27, 2010]. Garmin, 2009. [Online]. Available: http://www.garmin.com/.
6. Roccetti M, Marfia G, Amoroso A. An optimal 1D vehicular accident warning algorithm for realistic scenarios, In *Proceedings of the IEEE Symposium on Computers and Communications*, Riccione, Italy, 2010.
7. Elbes M, Al-Fuqaha A, Guizani M, Rayes A, Oh JS. A new hierarchical and adaptive protocol for minimum-delay V2V communication, In *Proceedings of the IEEE 2009 Global Telecommunications Conference*, Hololulu, Hawaii, USA, 2009.
8. Palazzi CE, Ferretti S, Roccetti M. An inter-vehicular communication architecture for safety and entertainment. *IEEE Transactions on Intelligent Transportation Systems* 2010; **11**(1): 90–99.
9. Sentinelli A, Celetto L, Lefol D, Palazzi CE, Pau G, Zahariadis T, Jari A. *Survey on P2P Overlay Streaming Clients, Towards the Future Internet*. IOS Press: Amsterdam, The Netherlands, 2009. 273-28.
10. Marfia G, Pau G, De Sena E, Giordano E, Gerla M. Evaluating vehicle network strategies for downtown portland: opportunistic infrastructure and the importance of realistic mobility models, In *Proceedings of the First ACM/SIGMOBILE International Workshop on Mobile Opportunistic Networking*, San Juan, Puerto Rico, USA, 2007.
11. Nzouonta J, Rajgure N, Guiling W, Borcea C. VANET routing on city roads using real-time vehicular traffic information. *IEEE Transaction on Vehicular Technology* 2009; **58**(7): 3609–3626.
12. Huang D, Shere S, Ahn S. Dynamic highway congestion detection prediction based on shock waves, In *Proceedings of the Seventh ACM International Workshop on Vehicular Internetworking*, Chicago, Illinois, USA, 2010.
13. Aman AJ, Nunez DC, Hamideddin M, Laungani P, Abbo W. Evaluation of congestion pricing for management highway in Seattle, In *Proceedings of the 2009 Systems and Information Engineering Design Symposium*, Charlottesville, Virginia, USA, 2009.
14. Palazzi CE. Residual capacity estimator for TCP on wired/wireless links, In *Proceedings of the WCC2004 Student Forum IFIP World Computer Congress*, Toulouse, France, 2004.
15. Kapoor R, Chen LJ, Lao L, Gerla M, Sanadidi MY. CapProbe: a simple and accurate capacity estimation technique, In *Proceedings of the 2004 Conference on Applications, Technologies, Architectures and Protocols for Computer Communications*, Portland, Oregon, USA, 2004.
16. Research and Innovative Technology Administration Bureau of Transportation Statistics. National Household Travel Survey Daily Travel Quick Facts, [Accessed: May 27, 2010]. RITA, May 2002. [Online]. Available: http://www.bts.gov/programs/national_household_travel_survey/daily_travel.html..
17. Wei CH, Lee Y. Development of freeway travel time forecasting models by integrating different sources of traffic data. *IEEE Transactions on Vehicular Technology* 2007; **56**(6): 3682–3694.
18. Bertini R. You are the traffic jam: an examination of congestion measures, department of civil and environmental engineering, Portland State University, 2005. Unpublished Research Monograph.
19. Lomax T, Turner S, Shunk G. Quantifying congestion. *Tech. Rep. 398*, Transportation Research Board, Washington D.C, US, 1997.
20. Yoon J, Noble B, Liu M. Surface Street Traffic Estimation, In *Proceedings of the Fifth ACM International Conference on Mobile Systems, Applications and Services*, San Juan, Puerto Rico, US, 2007.
21. Transportation Research Board Technical Staff. *Highway Capacity Manual*. Transportation Research Board: Washington D.C., US, 2000.
22. Official Google Traffic Blog Spot. [Online]. Available: http://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html, [Accessed February 21st 2011].
23. Marfia G, Roccetti M, Vehicular congestion modeling estimation for advanced traveler information systems, In *Proceedings of the 3rd IFIP/IEEE Wireless Days Conference*, Venice, Italy, 2010.
24. Dowling Associates, Sierra Research, Wiltech. Arterial speed study. *Tech. Rep. 04–006*, Southern California Association of Governments, Los Angeles, CA, 2005.

25. Los Angeles Department of Transportation Studies and Reports, [Accessed February 21st 2011]. [Online]. Available: http://www.ladot.lacity.org/studies_reports.htm.

## AUTHORS' BIOGRAPHIES

**Gustavo Marfia** received his Laurea degree in Telecommunications Engineering at the University of Pisa in 2003 and his PhD in Computer Science from the University of California, Los Angeles, in 2009. He is currently a postdoctoral researcher at the Department of Computer Science of the University of Bologna. His research interest includes transportation, biomedical, and entertainment systems.

**Marco Roccetti** received his Italian Laurea degree in Electronic Engineering from the University of Bologna, Bologna, Italy. He has been a full professor at the Computer Science Department, University of Bologna, since 2000. His research interests include digital audio and video, computer entertainment, and mobile Web applications, fields in which he has authored almost 200 technical refereed papers.

**Alessandro Amoroso** obtained his Laurea degree in Physics at the University of Bologna in 1987. He is currently an associate professor at the Department of Computer Science of the University of Bologna. His research interests include distributed systems, multimedia systems, and mobile devices.