

# A Max Coverage Formulation for Information Dissemination in Vehicular Networks

O. Trullols<sup>†</sup>, M. Fiore<sup>\*</sup>, C. Casetti<sup>\*</sup>, C.-F. Chiasserini<sup>\*</sup>, J.M. Barcelo-Ordinas<sup>†</sup>

<sup>†</sup> Universitat Politècnica de Catalunya, Barcelona, Spain

<sup>\*</sup> Dipartimento di Elettronica, Politecnico di Torino, Torino, Italy

**Abstract**—We consider that  $k$  infrastructured nodes have to be deployed for disseminating information to vehicles travelling in an urban area. We name the infrastructure nodes Dissemination Points (DPs) and formulate our problem as a Maximum Coverage Problem (MCP) so as to maximize the number of vehicles that get in contact with the DPs. Since the MCP is NP-hard, we solve it through heuristic algorithms, which present different levels of complexity and require different knowledge on the system. Then, we address the problem of covering a large number of vehicles for a sufficient amount of time. We therefore give a different formulation of the problem, which however is still NP-hard and requires a heuristic approach to be solved. Evaluation of the proposed solutions in a realistic urban environment shows how knowledge of vehicular mobility plays a major role in achieving an optimal coverage of mobile users, and that simple heuristics provide near-optimal results even in large-scale scenarios.

## I. INTRODUCTION

Vehicular Ad hoc NETWORKS (VANETs) are considered one of the most promising areas of scientific research in wireless networking. Undoubtedly, VANETs benefit from the cross-fertilization from the sprawling branches of ad hoc networks and from the wide range of safety, comfort and infotainment applications. Beside the obvious interest in safety aspects, a variety of non-safety applications of VANETs are being proposed and investigated. Information dissemination, lookup and retrieval, for a large number of purposes (parking lot availability, transportation timetables, pollution data collection), are at the forefront of non-safety-related VANET research. Also, the two main communication paradigms, namely vehicle-to-vehicle (V2V) and vehicles-to-infrastructure (V2I), provide different takes on the problem.

It should be pointed out that VANETs share, and possibly exacerbate, the typical shortcomings of ad hoc networks. Specifically: fleeting connectivity, rapidly shifting topologies, highly dynamic traffic patterns, constrained node movements. More related to the typical challenges in DTN (Disruption-Tolerant Networks) [1], [2] than in infrastructure-based wireless networks, the hurdles that VANETs have to overcome before being considered for widespread adoption are quite daunting. In this paper, we try to provide some answers to one of the recurring problems in V2I communications, i.e., that of information dissemination to passing vehicles. Assuming that an area, served by roads of an arbitrary given topology, must be equipped for information dissemination through the deployment of a limited number  $k$  of infrastructured nodes (e.g., IEEE 802.11 access points), what is the best deployment

strategy to maximize the dissemination of information?

In the following, we refer to the infrastructured nodes as *Dissemination Points* (DPs), and, as a first step, we show that road intersections are preferred locations to place DPs. Then, we address two different cases. Firstly, we assume that the information is just a small, self-contained item. A vehicle will receive the information item if it gets in contact with a DP at least once. Under this assumption, we are interested in placing the DPs at  $k$  of the possible intersections so as to maximize the number of vehicles that enter a DP coverage area at least once; we therefore model our problem as a Maximum Coverage Problem (MCP). Secondly, we consider the case in which vehicle-to-DP contact times have an impact on the dissemination process. In this case, we give a different formulation for our problem, which aims at favoring both the number of contacted vehicles as well as the contact times. Both versions of the problem, however, are NP-hard, thus we propose heuristic algorithms for their solution and evaluate the performance of the heuristics by considering a real-world urban environment and realistic vehicular traces.

## II. SYSTEM SCENARIO AND GOALS

We consider a urban road topology of area size equal to  $A$  and including  $N$  intersections. We assume that each DP has a dissemination range equal to  $R$ . Such a dissemination range may map into the DP's transmission range, or into its service range if dissemination can be performed through multihop communication. Also, we denote by  $V$  the number of vehicles that transit over the area  $A$  during a given time period, hereinafter called observation period.

Our goal is to deploy  $k$  DPs so as to maximize either the number of vehicles, among the possible  $V$ , served (i.e., covered) by the DPs, or to favor both the number of covered vehicles and the connection time between vehicles and DPs. This significantly differs from other coverage problems, since

- the DPs deployed in the area do not have to necessarily form a connected network or provide a continuous coverage of the road topology: this is one of the major differences from previous work on maximum graph coverage [3] as well as on cellular and sensor wireless networks (see e.g., [4]);
- vehicles may cross several intersections, thus they may be covered (i.e., served) by more than one DP. When contact time is taken into account, this aspect makes existing generalizations of the MCP unsuitable to our problem.

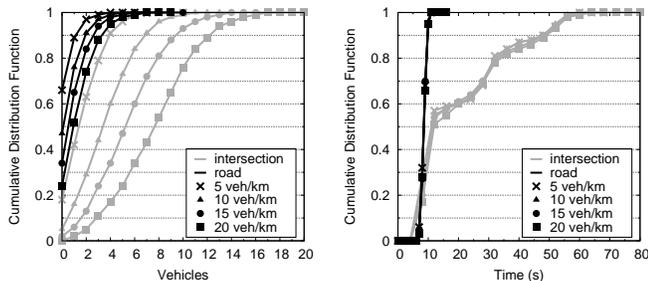


Fig. 1. CDF of the number of vehicles within range of the DP (left) and of the time spent by vehicles within range of the DP (right), with  $R = 50$  m and for different vehicle densities

### III. SELECTING THE LOCATION TYPE

The evaluation of *where* on a road to deploy the DPs is an important first step in designing an efficient dissemination system for vehicular environments. Nominally, the position of a DP over a single road segment can span anywhere between adjacent intersections: thus, the problem basically lies in deciding whether a DP should be located midway through the road segment, or closer to the intersections bounding it.

To this end, we simulate a realistic vehicular mobility over a simple road topology, and measure the potential for information dissemination of an individual DP, deployed at first in the intermediate point of a road segment, and then at an intersection ending the same street. The movement of vehicles is simulated with VanetMobiSim [5], employing the IDM-LC model, which reproduces car-to-car interactions, stopping, braking and acceleration phenomena in presence of traffic lights at road junctions, and overtakings, as observed in real world [6].

In particular, we considered different vehicular lane densities, ranging between 5 and 20 vehicles/km, which map to low and dense traffic conditions, respectively. The potential for dissemination is evaluated in terms of number of concurrent vehicle-to-DP contacts and of time spent by each vehicle within the dissemination range  $R$  of the DP: a higher number of vehicles, as well as longer permanence times, are indicative of a higher potential for information dissemination, as more users can receive larger portions of the content provided by the DP.

Fig. 1 depicts the Cumulative Density Function (CDF) of such two metrics, when the DP is positioned along the road or at the intersection, with varying vehicular densities. It can be observed that the car density has a negligible impact on the time that vehicles spend within DP's dissemination range, while it strongly impacts the number of vehicles in that same area. In both cases, however, deploying the DP at the intersection leads to better results, since more vehicles travel through the dissemination area, spending there a longer time.

We also analyzed the effect that different DP ranges have on the dissemination performance. Fig. 2 portrays the same metrics studied before, for several values of  $R$ . The dissemination range significantly affects both CDFs, with larger ranges clearly providing better performance. In any case, deploying

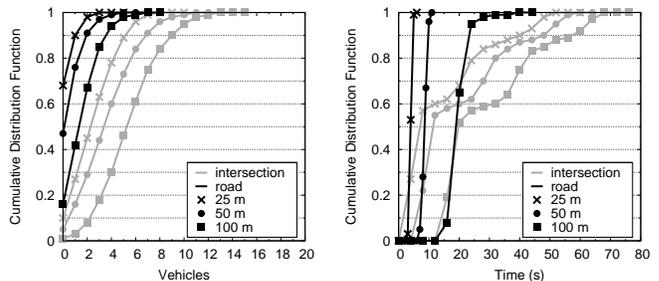


Fig. 2. CDF of the number of vehicles within range of the DP (left) and of the time spent by vehicles within range of the DP (right), with a vehicular lane density of 10 vehicles/km and for different DP dissemination ranges

the DP at the intersection yields again more favorable properties than positioning it along the road, for any value of  $R$ .

According to these results, intersections prove to be much better locations than road segments for the deployment of DPs, in terms of information dissemination potential. Thus, in the remainder of the paper, we will focus on the problem of DPs deployment at intersections of the road topology.

### IV. DEPLOYMENT ALGORITHMS

As stated before, we consider two cases, accounting for (i) only the number of vehicles that get in contact with DPs, and (ii) both the number of served vehicles and the vehicle-to-DP contact times.

#### A. Maximizing contacts

Our goal is to maximize the number of vehicles covered by  $k$  DPs. Based on the above results, we constrain ourselves to considering only the  $N$  intersections located in the road topology as possible locations for a DP. In particular, by analyzing the vehicular mobility in the selected area, we define an  $N \times V$  matrix  $\mathbf{P}$  whose generic element is given by

$$\mathbf{P}_{ij} = \begin{cases} 1 & \text{if vehicle } j \text{ crosses intersection } i \\ & \text{during the observation period} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

It is worth pointing out that the use of matrix  $\mathbf{P}$  requires that the identity of each vehicle be known so that it can be tracked across all intersections. (In Section IV-A3, we will relax this assumption and present an approach where the identity need not be recorded.)

We model the problem as a Maximum Coverage Problem (MCP), which can be formulated as follows. We are given a collection of sets  $\mathcal{S} = \{S_1, S_2, \dots, S_N\}$ , where each set  $S_i$  is a subset of a given ground set  $X = \{x_1, \dots, x_V\}$ . The goal is to pick  $k$  sets from  $\mathcal{S}$  to maximize the cardinality of their union.

To better understand the correspondence with our problem, consider that the elements in  $X$  are the vehicles that transit over the considered road topology during the observation period. Also, for  $i = 1, \dots, N$  we have

$$S_i = \{x_j \in X, j = 1, \dots, V : \mathbf{P}_{ij} = 1\} \quad (2)$$

i.e.,  $S_i$  includes all vehicles that cross intersection  $i$  at least once over the observation period. Thus, by solving the above problem, we obtain the set of  $k$  intersections where a DP should be placed so as to maximize the number of covered vehicles.

Unfortunately, the MCP problem is NP-hard; however, it is well known that the greedy heuristic achieves an approximation factor of  $1 - (1 - \frac{1}{m})^m$ , where  $m$  is the maximum cardinality of the sets in the optimization domain [7]. We report the greedy heuristic below.

1) *The greedy algorithm:* The greedy heuristic (hereinafter also called MCP-g) picks at each step a set (i.e., an intersection) maximizing the weight of the uncovered elements.

Let us introduce an auxiliary set  $G$ . Let  $G \subseteq \mathcal{S}$  be a collection of sets and  $W_i$  ( $i = 1, \dots, N$ ) be the number of elements covered by  $S_i$ , but not covered by any set in  $G$ . The steps of the greedy heuristic are reported in Algorithm 1.

---

**Algorithm 1** The MCP-g heuristic

---

**Require:**  $k, \mathbf{P}, \mathcal{S}$

- 1:  $G \leftarrow \emptyset, C \leftarrow 0, U \leftarrow \mathcal{S}$
  - 2:  $W_i = \sum_{j=1}^V \mathbf{P}_{ij}, i = 1, \dots, N$
  - 3: **repeat**
  - 4:   Select  $S_i \in U$  that maximizes  $W_i$
  - 5:    $G \leftarrow G \cup S_i$
  - 6:    $C \leftarrow C + 1$
  - 7:    $U \leftarrow U \setminus S_i$
  - 8:    $W_i = \sum_{\substack{j=1 \\ j: x_j \notin G}}^V \mathbf{P}_{ij}, i = 1, \dots, N$
  - 9: **until**  $C = k$  or  $U = \emptyset$
- 

Note that, although such algorithm provides a very good approximation of the optimal solution, it requires:

- i) global knowledge of the road topology and network system,
- ii) the identity of the vehicles which have crossed the  $N$  intersections during the observation period.

Below, we propose i) a hierarchical algorithm which reduces the computational complexity by applying the *divide et impera* approach, and ii) a different problem formulation where the knowledge of the vehicles identity is not needed.

2) *The subzone algorithm:* We superimpose an overlay grid with cells of arbitrary, equal size on our road topology. We name a cell as *subzone* and denote the number of subzones by  $B = 2^L$  (with  $L \in \mathbb{N}_1$ ). We define a hierarchical structure consisting of  $L + 1$  levels, such that, at the generic level  $l$  ( $l = 0, \dots, L$ ), the unit area includes  $2^{L-l}$  subzones.

We start by solving the maximum coverage problem in each subzone (i.e.,  $l = 0$ ), and we find the optimum location of  $k_0$  DPs in every overlay grid. Then, at each step  $l \geq 1$ , we divide the area of the grid into  $2^{L-l}$  subzones, each twice the size of a single subzone at the previous step, and we select  $k_l$  intersections among the ones that were chosen at step  $l - 1$ . We repeat the procedure till the subzone area coincides with the area of the overlay grid (i.e.,  $l = L$ ).

The subzone heuristic, hereinafter also called MCP-sz, is reported in Algorithm 2.

---

**Algorithm 2** The MCP-sz heuristic

---

**Require:**  $k, \mathbf{P}, \mathcal{S}, 1 < B = 2^L$

- 1:  $\mathcal{S}' \leftarrow \mathcal{S}$
  - 2: **for**  $l = 0$  to  $L$  **do**
  - 3:   Divide the road topology into  $2^{L-l}$  cells of equal size
  - 4:   **for**  $m = 1$  to  $2^{L-l}$  **do**
  - 5:     Solve the MCP in the  $m$ -th subzone, by taking  $\mathcal{S}'$  as input set and  $k_l$  as the number of DPs to deploy
  - 6:     Remove from  $\mathcal{S}'$  the unselected intersections
  - 7:      $m \leftarrow m + 1$
  - 8:   **end for**
  - 9:    $l \leftarrow l + 1$
  - 10: **end for**
- 

Note that the value of  $k_l$  can be set so as to limit the number of intersections selected within each subzone at step  $l$  ( $l = 0, \dots, L$ ). As an example, for  $k \ll N$ , we found that the algorithm can be efficiently run by fixing  $k_l = k, \forall l$ . For larger values of  $k$ , instead, setting  $k_l = \lceil \frac{k}{2^{L-l}} \rceil + 2^{L-l} - 1$  allows the selection of at least  $\frac{k}{2^{L-l}}$  per subzone, i.e.,  $k$  intersections in the whole area, plus some extra intersections per subzone ( $2^{L-l} - 1$ ). The benefit of such redundancy is twofold: it allows us to better approximate a centralized solution, and its impact is limited since the number of extra intersections reduces exponentially at each step of the procedure till it reaches 0 at the last round (i.e.,  $l = L$ ).

As a last remark, the value of  $B$  can be determined so as to limit the number of candidate intersections that are selected at each round (hierarchical level) of the procedure. In particular, given  $k_0$ , the number of intersections selected in the first round ( $l = 0$ ) must be less than or equal to the number of existing intersections, i.e.,

$$Bk_0 \leq N \quad (3)$$

Since  $B = 2^L$ , from (3), it is possible to derive a value for  $L$  and, thus, for the number of levels that avoids useless iterations, i.e., to consider too fine grids which do not yield any selection of intersections.

3) *Unknown vehicles identity:* Here we assume that the vehicles identity is not recorded and the only available information is the number of different vehicles that have crossed each of the  $N$  intersections during the observation period. Thus, our objective becomes the maximization of the total number of service opportunities provided by  $k$  DPs.

To this end, let  $\nu_i, i = 1, \dots, N$ , be the total number of vehicles that have crossed intersection  $i$  during the observation period, i.e.,

$$\nu_i = \sum_{j=1}^V \mathbf{P}_{ij} \quad i = 1, \dots, N \quad (4)$$

We then model the problem as a 0-1 Knapsack Problem (KP), which is defined as follows [8]. We are given a bag and a set

of  $N$  items  $\mathcal{I} = \{I_1, \dots, I_N\}$ . Each item  $I_i \in \mathcal{I}$  has a non-negative value and a non-negative weight, and the maximum weight that we can carry in the bag is equal to  $k$ . The objective is to select a subset of items  $\mathcal{I}' \subseteq \mathcal{I}$  whose weight does not exceed  $k$  and that maximizes the overall value of the bag. Each item must only be selected once.

To better understand the correspondence with our problem, consider that the elements in  $\mathcal{I}$  are the intersections; each intersection  $i$  has a weight equal to 1 and a value equal to  $\nu_i$  ( $i = 1, \dots, N$ ). Thus, our problem can be formulated as,

$$\max \quad \sum_{i=1}^N \nu_i y_i \quad (5)$$

$$s.t. \quad \sum_{i=1}^N y_i \leq k; \quad y_i \in \{0, 1\} \quad \forall i \quad (6)$$

The 0-1 KP is an NP-hard problem in general, however in our case, where all intersections have the same weight, it can be solved in polynomial time by simply sorting the intersections in decreasing order by their value, and selecting the first  $k$  intersections. We will refer to this algorithm as KP-P.

In Section V-B, we present the deployment and coverage performance obtained by solving the MCP by brute force and through the greedy algorithm (MCP-g), compared against the cases where the hierarchical approach is used (MCP-sz) and where vehicles identity are not available (KP-P).

### B. Maximum coverage and contact times

Here we address our second case, where  $k$  DPs have to be deployed at the road intersections so as to favor both the number of covered vehicles, as well as the time for which they are covered. To this end, let us define an  $N \times V$  matrix  $\mathbf{T}$  whose generic element,  $\mathbf{T}_{ij}$  represents the total time that vehicle  $j$  would spend under the coverage of a DP if the DP were located at intersection  $i$ , i.e., the contact time between a vehicle  $j$  and a DP located at intersection  $i$ . Then, we formulate the following problem, which we name Maximum Coverage with Time Threshold Problem (MCTTP): given  $k$  DPs to be deployed, we aim at serving as many vehicles as possible, for (possibly) at least  $\tau$  seconds each, i.e.,

$$\max \quad \sum_{j=1}^V \left[ \min \left( \tau, \sum_{i=1}^N \mathbf{T}_{ij} y_i \right) \right] \quad (7)$$

$$s.t. \quad \sum_{i=1}^N y_i \leq k; \quad y_i \in \{0, 1\} \quad \forall i \quad (8)$$

Note that in (7) we place a DP at an intersection so as to maximize the number of vehicles that are covered, taking into account a vehicle's contact time up to a maximum value equal to  $\tau$ : DPs that provide coverage for at least  $\tau$  seconds to a given vehicle do not further contribute to the overall gain of covering such a vehicle. The constraint in (8) instead limits the number of DPs to  $k$ .

It can be easily verified that the MCP is a particular case of the above formulation, obtained by setting  $\tau = 1$  and  $\mathbf{T}_{ij} =$

$\mathbf{P}_{ij}$ . Hence, MCTTP is NP-hard and we propose the following heuristic for its solution.

1) *A greedy approach:* The greedy algorithm we propose to solve the MCTTP problem, denoted by MCTTP-g, picks an intersection at each step so as to maximize the provided coverage time, although only the contribution due to vehicles for which the threshold  $\tau$  has not been reached is considered.

Let  $G \subseteq \mathcal{S}$  be a collection of sets and let now  $W_i$  ( $i = 1, \dots, N$ ) be the total contact time provided by intersection  $i$ , considering for each vehicle a contribution such that the vehicle's coverage time due to  $G \cup S_i$  does not exceed the threshold  $\tau$ . The greedy heuristic is reported in Algorithm 3.

---

#### Algorithm 3 The MCTTP-g heuristic

---

**Require:**  $k, \mathbf{T}, \tau, \mathcal{S}$

- 1:  $G \leftarrow \emptyset, C \leftarrow 0, U \leftarrow \mathcal{S}$
  - 2:  $t_j = 0, j = 1, \dots, V$
  - 3: **repeat**
  - 4:  $W_i = \sum_{j=1}^V \min(\tau - t_j, \mathbf{T}_{ij}), i = 1, \dots, N$
  - 5: Select  $S_i \in U$  that maximizes  $W_i$
  - 6:  $G \leftarrow G \cup S_i$
  - 7:  $C \leftarrow C + 1$
  - 8:  $U \leftarrow U \setminus S_i$
  - 9:  $t_j = \min(\tau, t_j + \mathbf{T}_{ij}), j = 1, \dots, V$
  - 10: **until**  $C = k$  or  $U = \emptyset$
- 

Again, we notice that the time-threshold heuristic requires knowledge of the global road topology and of the vehicles identity. Likewise for the MCP, we present a time-subzone algorithm, which adopts the *divide et impera* approach and a 0-1 KP, for which knowledge of the vehicles' identity is not necessary.

2) *The time-subzone algorithm:* As done in Section IV-A2, we divide the road topology in  $B = 2^L$  cells, called subzones, and we apply the time-subzone heuristic (MCTTP-sz) whose steps are reported in Algorithm 4.

---

#### Algorithm 4 The MCTTP-sz heuristic

---

**Require:**  $k, \mathbf{T}, \mathcal{S}, 1 < B = 2^L$

- 1:  $\mathcal{S}' \leftarrow \mathcal{S}$
  - 2: Divide the road topology in  $B$  cells of equal size
  - 3: **for**  $l = 0$  to  $L$  **do**
  - 4:   **for**  $m = 1$  to  $2^{L-l}$  **do**
  - 5:     Solve the MCTTP in the  $m$ -th subzone, by taking  $\mathcal{S}'$  as input set and  $k_l$  as the number of DPs to deploy
  - 6:     Remove from  $\mathcal{S}'$  the unselected intersections
  - 7:      $m \leftarrow m + 1$
  - 8:   **end for**
  - 9:    $l \leftarrow l + 1$
  - 10: Merge each pair of adjacent subzones so as to obtain  $2^{L-l}$  subzones
  - 11: **end for**
-

3) *Unknown vehicles identity*: When the vehicles' identities are not available, the only information we have is the total time that all vehicles would spend under the coverage of a DP if it were located at intersection  $i$ , i.e.,

$$T_i = \sum_{j=1}^V \mathbf{T}_{ij} \quad i = 1, \dots, N \quad (9)$$

Thus, in this case we want to maximize the total contact (service) time offered to the vehicles, when  $k$  DPs are deployed. Again, the problem can be formulated as the following 0-1 KP. We are given a set of  $N$  intersections (items)  $\mathcal{I} = \{I_1, \dots, I_N\}$ ; each intersection has a value  $T_i$  and unitary weight, and the maximum number of selected intersections (maximum weight) must be equal to  $k$ . The objective is to select a subset of  $k$  intersections that maximizes the overall service time provided to the vehicles, i.e.,

$$\max \quad \sum_{i=1}^N T_i y_i \quad (10)$$

$$\text{s.t.} \quad \sum_{i=1}^N y_i \leq k; \quad y_i \in \{0, 1\} \quad \forall i \quad (11)$$

As already mentioned, the above problem can be solved in polynomial time by using the simple algorithm reported in Section IV-A3. We refer to this solution, which requires the knowledge of the  $T_i$  coefficients ( $i = 1, \dots, N$ ), as KP-T.

The performance of the brute force solution of the MCTTP problem are presented in Section V-C, together with those of its greedy (MCTTP-g), subzone (MCTTP-sz), and no-identity (KP-T) heuristics.

## V. PERFORMANCE EVALUATION

We applied the algorithms presented in the previous sections to a real-world road topology, in presence of realistic vehicular mobility. The resulting DP deployments were then evaluated in terms of information dissemination capabilities.

In this section, we first introduce the evaluation scenario, and then we compare the results obtained with the different deployment algorithms maximizing contacts, as well as the results obtained with the different algorithms maximizing coverage and contact times.

### A. Scenario

For our performance evaluation, we selected a real-world road topology from the canton of Zurich, in Switzerland. Realistic traces of the vehicular mobility in such region are available from the Simulation and Modelling Group at ETH Zurich [9]. These traces describe the individual movement of cars through a queue-based model calibrated on real data [10]: they thus provide a realistic representation of vehicular mobility at both microscopic and macroscopic levels.

The road topology we considered is a 100 km<sup>2</sup> portion of the urban and suburban area of the city of Zurich, a map of which can be found in the plots of Fig. 5. We extracted an hour and a half of vehicular mobility, in presence of average

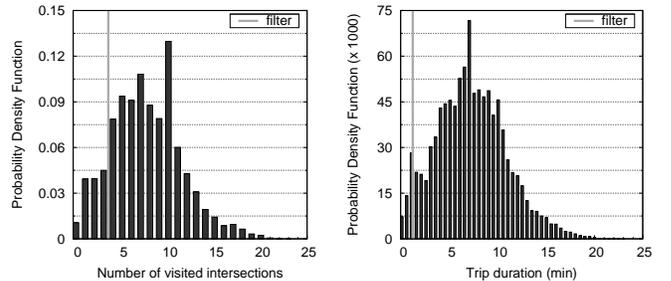


Fig. 3. PDF of the number of intersections traversed (left) and of the trip duration (right) for all vehicles in the trace

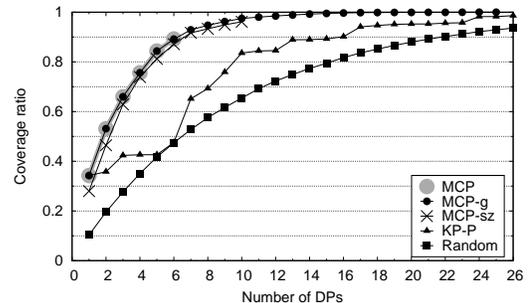


Fig. 4. Ratio of vehicles experiencing at least one contact with a DP versus the number  $k$  of DPs deployed, for different algorithms that maximize contacts

traffic density conditions, during which more than 70000 cars travel within the selected area.

In order to remove partial trips (i.e., vehicular movements starting or ending close to the border of the square area), we filtered the trace, by removing cars that traverse only three intersections or less, as well as those spending less than one minute in the considered region. Fig. 3 shows where the filtering thresholds fall with respect to the probability density functions (PDF) of visited intersections and trip duration: we can notice that the selected thresholds result in a low percentage of cars being removed from the trace.

### B. Maximizing contacts

In the scenario described above, we first run the deployment algorithms for contact maximization presented in Section IV-A. The selected settings for the MCP-sz algorithm were  $L=4$ ,  $k_l=k$ . Also, in order to provide a lower-bound benchmark to the performance of the schemes, we also test a random deployment, that ignores the vehicular mobility information and whose outcome results from averaging multiple tests over the same road topology.

The coverage ratio, i.e., the number of vehicles that experience at least one contact with a DP over the total number of vehicles in the scenario, is shown in Fig. 4. For each deployment algorithm, the ratio is recorded versus the number of allowed DPs  $k$ .

Three different behaviors can be distinguished. The first is that of the random algorithm, which, lacking all information on the movement of vehicles, performs poorly: it cannot provide one contact or more per vehicle even when a large number of

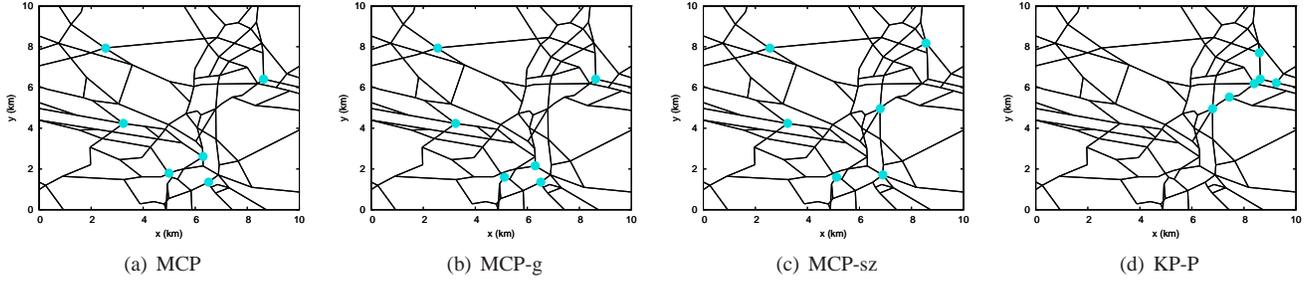


Fig. 5. Deployments of DPs obtained with different algorithms maximizing contacts over the Zurich road topology, for  $k = 6$

DPs is deployed (we stress that 26 DPs occupy roughly one third of the intersections in the road topology under study).

The second behavior is that of the KP scheme, which has only partial knowledge of the vehicular mobility, since it accounts for vehicular densities at intersections but neglects the mobility in between them. The KP algorithm performs better than the random one, although its absolute result still has wide margins for improvement. As a matter of fact, the curve gets close but does not reach a coverage ratio equal to one, even when 26 DPs are employed. Moreover, the progress in terms of covered vehicles is quite irregular as the number  $k$  of deployed DPs grows. This suggests that the KP scheme can at times pick new DPs that do not improve the overall coverage.

The third behavior is that shown by the remaining algorithms: the brute force solution to the MCP, the greedy solution, and the subzone solution. The common point to these algorithms is that they all exploit full knowledge of the vehicles identity and mobility over the road topology. It is interesting to notice how both the greedy and the subzone schemes almost overlap with the optimal solution, and thus provide an excellent result in terms of information dissemination. Also, we stress that the difference with respect to the KP algorithm is extremely high, since the greedy and subzone schemes cover 90% of vehicles with just 6 DPs (accounting for 7% of the available intersections), and 99% of vehicles with 12 DPs (14% of intersections).

Further insight in the different behaviors is provided in Fig. 5. The figure shows the actual positions of the DPs over the road topology, when  $k = 6$ , for the MCP, MCP-g, MCP-sz, and KP-P formulations. There, it can be observed how the greedy algorithm results in a solution that is nearly identical to the optimal one, whereas the subzone solution is less similar to the optimal, but still close to it. The reason is that the hierarchical approach trades the reduction in complexity for optimality, and can take suboptimal decisions during initial iterations. However, the final result is still very close to that obtained by solving the MCP by brute force. On the contrary, the deployment achieved by the KP algorithm is noticeably different, as DPs tend to be gathered in a same area, characterized by high vehicular traffic density. Since the selected intersections are close to each other, a high number of vehicles travels through several of the deployed DPs, so that most of the DPs have a very small impact on the coverage.

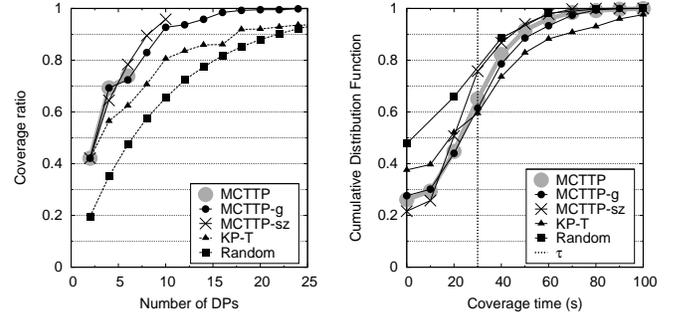


Fig. 6. Coverage ratio versus the number  $k$  of DPs deployed, for  $\tau=30s$  (left) and coverage time CDF, for  $\tau=30s$  and  $k=6$  (right)

By summarizing the results, we can conclude that:

- 1) knowledge of vehicular trajectories is the discriminating factor in achieving an optimal deployment of DPs;
- 2) when exploiting such a knowledge, even a simple, hierarchical solution, such as the subzone algorithm can lead to near-optimal results in real-world road topologies of tens of  $km^2$ ;
- 3) exploiting these properties it is possible to inform a high percentage of vehicles by deploying DPs at a small percentage of intersections.

### C. Maximizing coverage and contact times

Taking into account the time dimension, we increase the complexity of the problem, by maximizing coverage and contact times between vehicles and DPs. A random deployment is again employed to benchmark the performance of the algorithms we introduced in Section IV-B.

The coverage ratio achieved by the different schemes as the number of deployed DPs  $k$  varies is depicted in the left plot of Fig. 6, for a time threshold value  $\tau = 30s$ . Exactly as observed in the previous section, also in this case information on vehicular mobility plays a major role in favoring contacts among vehicles and DPs. As a matter of fact, the random solution performs poorly, while the KP-T algorithm provides a better coverage of vehicles. The MCTTP, MCTTP-g, and MCTTP-sz solutions, leveraging their knowledge of cars trajectories, guarantee the highest coverage and tend to perform similarly.

Fig. 6 also reports, in the right plot, the distribution of the coverage time, i.e., the amount of time that each vehicle spends within range of DPs during its trip in the considered scenario. The figure refers to the case in which 6 DPs are deployed

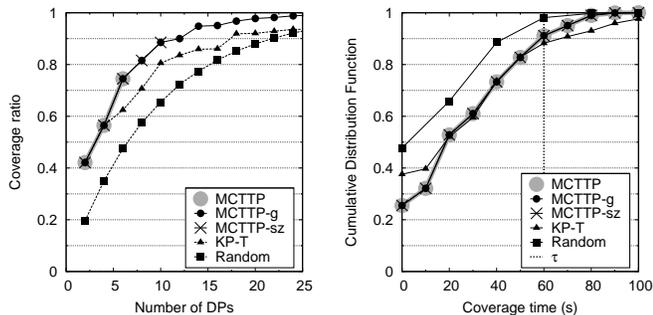


Fig. 7. Coverage ratio versus the number  $k$  of DPs deployed, for  $\tau=60s$  (left) and coverage time CDF, for  $\tau=60s$  and  $k=6$  (right)

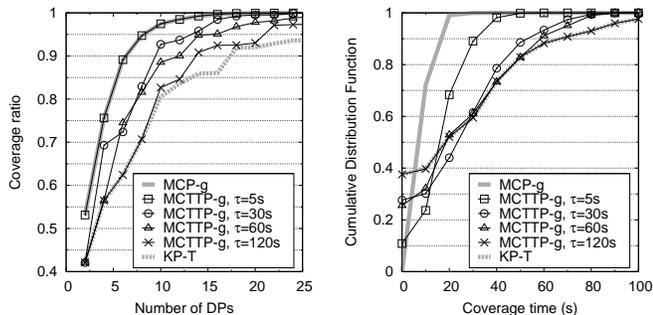


Fig. 8. Coverage ratio versus the number  $k$  of DPs deployed (left) and coverage time CDF, for  $k=6$  (right)

over the road topology, and  $\tau$  is set to 30s, as remarked by the vertical threshold line in the plot. The time coverage goal is to maximize the number of vehicles that spend  $\tau$  seconds or more under coverage of DPs: it translates into the CDF curve crossing  $\tau$  at the lowest possible point.

When comparing the two plots in Fig. 6, we can notice that MCTTP and MCTTP-g provide a matching performance, which is superior to those achieved by the other schemes. Indeed, a random deployment of DPs induces both a lower number of vehicle-to-DP contacts and a shorter coverage time with respect to MCTTP and MCTTP-g. The KP-T solution leads to a slightly better performance in terms of coverage time (2% more cars covered for at least  $\tau$  seconds), which is however paid at a high coverage ratio cost (more than a 10% reduction in covered cars). The MCTTP-sz algorithm instead allows a 5% increase in car-to-DP contacts, but the 30s coverage threshold is reached by 15% less vehicles.

Similar behaviors can be observed also for different values of the  $\tau$  threshold. In Fig. 7,  $\tau$  is set to 60s, and we can observe that the random and KP-T schemes still perform worse than MCTTP and its greedy and subzone heuristics.

The relationship of the  $\tau$ -dependent schemes from  $\tau$  is studied in Fig. 8. There, we focus on MCTTP-g, since the other algorithms showed similar behaviors, and evaluate it as  $\tau$  ranges between 5 and 120 seconds. The coverage ratio, in the left plot of Fig. 8, shows how the MCTTP-g solution falls in between those obtained with an algorithm that maximizes vehicle-to-DP contacts, i.e., MCP-g, and with one that maximizes the overall coverage time, i.e., KP-T. In particular, for low values of  $\tau$ , MCTTP-g tends to MCP-g, since the time

constraint is easily satisfied (a contact with a single DP is often sufficient to reach the desired coverage time) and the algorithm can thus focus on maximizing the coverage. On the other hand, when  $\tau$  is high, MCTTP-g tends to KP-T, since the desired coverage time is seldom reached, and thus the same vehicles keep on contributing to the optimization: the focus of the algorithm then shifts onto coverage times.

This is confirmed by the coverage time CDFs, on the right plot of Fig. 8, where the same behavior of the MCTTP-g algorithm is observed, as  $\tau$  varies. It can be however noted how MCTTP-g with  $\tau=5s$  matches MCP-g in terms of coverage ratio, but outperforms it in terms of coverage time. Similarly, MCTTP-g with  $\tau=120s$  matches KP-T as far as the coverage time is concerned, but provides a better coverage ratio. The combined maximization of contacts and coverage time can thus achieve better performance than contacts-only or time-only driven solutions even in borderline conditions.

Summarizing, we can draw the following conclusions:

- also when maximizing contacts and coverage time, simple, hierarchical solutions that exploit knowledge of vehicular mobility can lead to quasi-optimal results over large-scale road topologies;
- the coverage time threshold  $\tau$  can be used to calibrate the deployment so that it is preferably driven by vehicle-to-DP contacts or by coverage time.

## VI. CONCLUSIONS

We proposed a maximum coverage approach to the problem of information dissemination in vehicular networks. The formulations and relative heuristics we presented tackle both the case in which maximizing vehicle-to-DP contacts is the only goal, as well as the case in which coverage time is also an important aspect to account for. We evaluated the different solutions in a real world topology, showing that knowledge of vehicular mobility is the main factor in achieving an optimal deployment of DPs. Our results also prove that, given such knowledge, simple heuristics can be successfully employed to plan a deployment capable of informing more than 95% of vehicles with a few DPs.

## REFERENCES

- [1] X. Zhang, J. Kurose, B. Levine, D. Towsley, H. Zhang, "Study of a Bus-based Disruption-tolerant Network: Mobility Modeling and Impact on Routing," *ACM MobiCom*, Montreal, Canada, 2007.
- [2] J. Burgess, B. Gallagher, D. Jensen, B. N. Levine, "Maxprop: Routing for Vehicle-based Disruption-tolerant Networks," *IEEE INFOCOM*, Barcelona, Spain, 2006.
- [3] S. O. Krumke, M. V. Marathe, D. Poensgen, S. S. Ravi, H.-C. Wirth, "Budgeted Maximum Graph Coverage," *Lecture Notes In Computer Science*, vol. 2573, pp. 321–332, 2002.
- [4] R. M. Whitaker, L. Raisanen, S. Hurley, "The Infrastructure Efficiency of Cellular Wireless Networks," *Computer Networks*, vol. 48, no. 6, pp. 941–959, 2005.
- [5] VanetMobiSim, <http://vanet.eurecom.fr>.
- [6] M. Fiore, J. Hari, F. Filali, C. Bonnet, "Vehicular Mobility Simulation for VANETs," *IEEE/SCS Annual Simulation Symposium*, 2006.
- [7] A. A. Ageev, M. I. Sviridenko, "Approximation Algorithms for Maximum Coverage and Max Cut with Given Sizes of Parts," *Lecture Notes in Computer Science*, vol. 16, 1999.
- [8] D. Pisinger, "Where Are the Hard Knapsack Problems?," *Technical Report 2003/08*, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark, 2003.
- [9] ETH traces, <http://lst.inf.ethz.ch/ad-hoc/car-traces>.
- [10] N. Cetin, A. Burri, K. Nagel, "A Large-scale Multi-agent Traffic Microsimulation Based on Queue Model," *STRC'03*, Ascona, Switzerland, 2003.