Computer Communications 33 (2010) 432–442

Contents lists available at ScienceDirect

Computer Communications

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

Planning roadside infrastructure for information dissemination in intelligent transportation systems

O. Trullols [a], M. Fiore [b], C. Casetti [c], C.F. Chiasserini [c,*], J.M. Barcelo Ordinas [a]

[a] *Departament d'Arquitectura de Computadors, Universitat Politècnica de Catalunya, C/ Jordi Girona 1-3, Barcelona, Spain*
[b] *Université de Lyon, INSA Lyon, INRIA, CITI, F-69621, France*
[c] *Dipartimento di Elettronica, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy*

ARTICLE INFO

Article history:
Available online 4 December 2009

Keywords:
Vehicular networks
Network deployment
Maximum coverage

ABSTRACT

We consider an intelligent transportation system where a given number of infrastructured nodes (called Dissemination Points, DPs) have to be deployed for disseminating information to vehicles in an urban area. We formulate our problem as a Maximum Coverage Problem (MCP) and we seek to maximize the number of vehicles that get in contact with the DPs over the considered area. The MCP is known to be NP-hard in its standard formulation, therefore we tackle it through heuristic algorithms, which present different levels of complexity and require different knowledge on the system. Next, we address the problem of guaranteeing that a large number of vehicles travel under the coverage of one or more DPs 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. By evaluating the proposed solutions in a realistic urban environment, we observe that simple heuristics provide near-optimal results even in large-scale scenarios. However, we remark that a near-optimal coverage of mobile users can be achieved only when the characteristics of vehicular mobility are known.

1. Introduction

Wireless communications for intelligent transportation systems (ITS) are intended for the support of traffic safety and efficiency, as well as of value-added services, such as infotainment and commercial applications. The main components of the ITS architecture are roadside infrastructures and vehicles, and two main communication paradigms are foreseen, namely vehicle-to-vehicle (V2V) and vehicles-to-infrastructure (V2I). As a consequence, most of the research efforts so far have focused on the development of protocols and applications suitable for an ad hoc network composed of vehicles (the so-called VANET) and infrastructure nodes.

In this work, we deal with information dissemination to passing vehicles, tackling the specific issue of deploying an intelligent transportation infrastructure that efficiently achieves the dissemination goal. More specifically, we consider a system that has to support information dissemination or lookup and retrieval, for such purposes as parking lot availability, transportation timetables, pollution data collection. Then we pose the following question: assuming that an area, with an arbitrary road topology, must be equipped with 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?

It should be pointed out that vehicular networks share, and possibly exacerbate, the typical shortcomings of ad hoc networks. Specifically: fleeting connectivity, rapidly shifting topologies, highly dynamic traffic patterns, constrained node movements. In particular, unlike cellular communication networks, vehicular networks do not necessarily need continuous coverage, rather, they can be supported by hot spots in correspondence of roadside infrastructure nodes, which provide intermittent connectivity to vehicles. The challenges featured by this scenario are therefore more related to the ones typically found in DTN (Disruption-Tolerant Networks) [1,2] than in infrastructure-based wireless networks.

In principle, an information dissemination system could leverage both V2V and V2I communications: when only few road-side units can be deployed, V2V communications enable data sharing thus increasing the throughput perceived by the users while downloading a content. However, the gain achieved through V2V communications strictly depends on the particular cooperation paradigm adopted for the dissemination, and it is thus difficult to evaluate in the general case. In this work, we start by studying the problem of optimally positioning infrastructure nodes only considering V2I communications: when dealing with schemes that also exploit V2V communications, our approach results in a generic worst-case analysis.

* Corresponding author.
E-mail addresses: trullols@ac.upc.edu (O. Trullols), marco.fiore@insa-lyon.fr (M. Fiore), casetti@tlc.polito.it (C. Casetti), chiasserini@tlc.polito.it (C.F. Chiasserini), joseb@ac.upc.edu (J.M. Barcelo Ordinas).

0140-3664/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2009.11.021

We refer to the infrastructured nodes as *Dissemination Points* (DPs), and, as a first step to the solution of our problem, 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. Note that other performance metrics and, thus, optimization objectives could be also considered (e.g., minimizing the information dissemination time), but, again, it would require further assumptions on the knowledge that is possible to acquire (e.g., on content size, per-vehicle link data rate, etc.).

The performance of our heuristics is evaluated by considering a real-world urban environment and realistic vehicular traces. More specifically, we use traces of vehicular mobility in the canton of Zurich that have a duration of one hour and a half [3]. We point out, however, that in presence of very long traces, our models and solutions should be applied to rush-hour representations of the vehicular traffic, as typically done in system planning.

The remainder of the paper is organized as follows. Section 2 introduces the network scenario under study, while Section 3 justifies the choice of intersections as best locations to deploy DPs. The DP deployment problem is formulated as an optimization problem in Section 4, both when the number of contacts and when the contact time are considered as target performance metrics. Results derived through realistic simulations and vehicular traces are presented in Section 5. Finally, Sections 6 and 7, respectively, review previous work and draw some concluding remarks.

2. System scenario and goals

We consider a urban road topology of area size equal to A, 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. Note that 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; also, energy saving is not one of our goals. These are major differences with respect to previous work on maximum graph coverage [4] as well as on cellular and sensor wireless networks (see e.g., [5–8]);
- vehicles directly access the DPs; in addition their movement obeys traffic regulations and is constrained by the road topology; as a consequence, the scenario differs from the one studied in [9] for the deployment of Internet access points in static networks, or from mobile sensor networks [10];
- 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.

In the following, we deal with the problem of planning vehicular networks for information dissemination, taking into account the above issues and the peculiarities of these systems.

3. 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 [11], 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 overtaking, as observed in real world [12].

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 lingering 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 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.

4. 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.

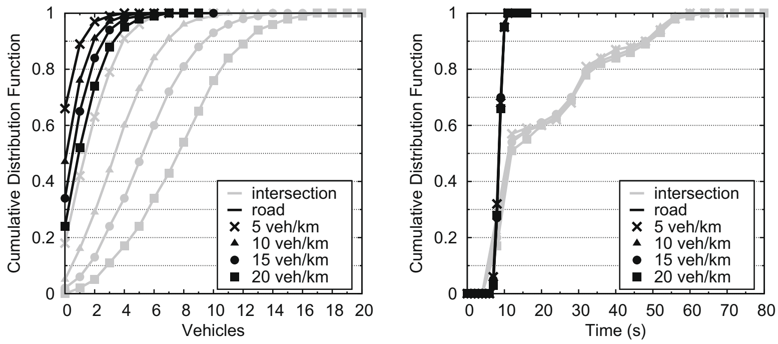


Fig. 1. CDF of the number of vehicles within range of the DP (left) and of the time spent by vehicles within range of the DP (right), with $R = 50$ m and for different vehicle densities.

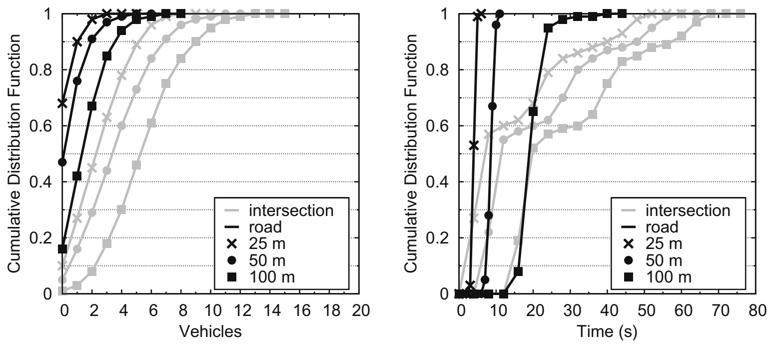


Fig. 2. CDF of the number of vehicles within range of the DP (left) and of the time spent by vehicles within range of the DP (right), with a vehicular lane density of 10 vehicles/km and for different DP dissemination ranges.

4.1. 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 4.1.3, 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 $\mathscr{S} = \{S_1, S_2, \ldots, S_N\}$, where each set S_i is a subset of a given ground set $X = \{x_1, \ldots, x_V\}$. The goal is to pick k sets from $\mathscr{S}$ 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 $i = 1, \ldots, N$ we have

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

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 [13]. We report the greedy heuristic below.

4.1.1. 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 $G \subseteq \mathscr{S}$ be a collection of sets and W_i $(i = 1, \ldots, N)$ 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.

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.

Algorithm 1. The MCP-g heuristic

```
Require: k, P, 𝒮
  G ← ∅, C ← 0, U ← 𝒮
  W_i = Σ_{j=1}^{V} P_ij, i = 1,...,N
  repeat
    Select S_i ∈ U that maximizes W_i
    G ← G ∪ S_i
    C ← C + 1
    U ← U \ S_i
    W_i = Σ_{j=1, j: x_j ∉ G}^{V} P_ij, i = 1,...,N
  until C = k or U = ∅
```

4.1.2. 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 ($l = 0,\ldots,L$), 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 $l \geqslant 1$, 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, P, 𝒮, 1 < B = 2^L
  𝒮' ← 𝒮
  for l = 0 to L do
    Divide the road topology into 2^(L−l) cells of equal
    size
    for m = 1 to 2^(L−l) do
      Solve the MCP in the m-th subzone, by taking
      𝒮' as input set and k_l as the number of DPs to
      deploy
      Remove from 𝒮' the unselected intersections
      m ← m + 1
    end for
    l ← l + 1
  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 ($l = 0,\ldots,L$). 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.,

$$Bk_0 \leqslant N \tag{3}$$

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.

4.1.3. Unknown vehicles identity

Unlike the previous case, we now 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 $v_i,\ i = 1,\ldots,N$, be the total number of vehicles that have crossed intersection i during the observation period, i.e.,

$$v_i = \sum_{j=1}^{V} \mathbf{P}_{ij} \quad i = 1,\ldots,N \tag{4}$$

We then model the problem as a 0–1 Knapsack Problem (KP), which is defined as follows [14]. We are given a bag and a set of N items $\mathscr{I} = \{I_1,\ldots,I_N\}$. Each item $I_i \in \mathscr{I}$ 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 $\mathscr{I}' \subseteq \mathscr{I}$ 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 I are the intersections; each intersection i has a weight equal to 1 and a value equal to v_i ($i = 1,\ldots,N$). Thus, our problem can be formulated as

$$\max \quad \sum_{i=1}^{N} v_i y_i \tag{5}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} y_i \leqslant k; \quad y_i \in \{0,1\} \forall i \tag{6}$$

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 5.2, 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).

4.2. 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 τ seconds each, i.e.,

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

$$\text{s.t} \quad \sum_{i=1}^{N} y_i \leqslant k; \quad y_i \in \{0,1\} \forall i \tag{8}$$

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 τ: DPs that provide coverage for at least τ 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.

4.2.1. 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 τ has not been reached is considered.

Let $G \subseteq \mathcal{S}$ be a collection of sets and let now W_i $(i = 1, \ldots, N)$ 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 τ. The greedy heuristic is reported in Algorithm 3.

Algorithm 3. The MCTTP-g heuristic

Require: $k, \mathbf{T}, \tau, \mathcal{S}$
- $G \leftarrow \emptyset, C \leftarrow 0, U \leftarrow \mathcal{S}$
- $t_j = 0, j = 1, \ldots, V$
- **repeat**
  - $W_i = \sum_{j=1}^{V} \min(\tau - t_j, \mathbf{T}_{ij}), i = 1, \ldots, N$
  - Select $S_i \in U$ that maximizes W_i
  - $G \leftarrow G \cup S_i$
  - $C \leftarrow C + 1$
  - $U \leftarrow U \setminus S_i$
  - $t_j = \min(\tau, t_j + \mathbf{T}_{ij}), j = 1, \ldots, V$
- **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.

4.2.2. The time-subzone algorithm

As done in Section 4.1.2, 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.

4.2.3. 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.,

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

Algorithm 4. The MCTTP-sz heuristic

Require: $k, \mathbf{T}, \mathcal{S}, 1 < B = 2^L$
- $\mathcal{S}' \leftarrow \mathcal{S}$
- Divide the road topology in B cells of equal size
- **for** $l = 0$ to L **do**
  - **for** $m = 1$ to 2^{L-l} **do**
    - Solve the MCTTP in the mth subzone, by taking $\mathcal{S}'$ as input set and k_l as the number of DPs to deploy
    - Remove from $\mathcal{S}'$ the unselected intersections
    - $m \leftarrow m + 1$
  - **end for**
  - $l \leftarrow l + 1$
  - Merge each pair of adjacent subzones so as to obtain 2^{L-l} subzones
- **end for**

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 a 0–1 KP. We are given a set of N intersections (items) $\mathcal{I} = \{I_1, \ldots, I_N\}$; 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.,

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

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

As already mentioned, the above problem can be solved in polynomial time by using the simple algorithm reported in Section 4.1.3. We refer to this solution, which requires the knowledge of the T_i coefficients $(i = 1, \ldots, N)$, as KP-T.

The performance of the brute force solution of the MCTTP problem are presented in Section 5.3, together with those of its greedy (MCTTP-g), subzone (MCTTP-sz), and no-identity (KP-T) heuristics.

4.3. Computational complexity

The computational complexity of both MCP and MCTTP is $O(VN^k)$: given N intersections, all possible combinations where the k DPs can be placed have to be considered and the weight of each intersection is computed over V vehicles. The cost of both greedy heuristics, MCP-g and MCTTP-g, is $O(kVN)$, since, for k times, we have to select the best choice among the candidate intersections (initially set to N), and again the selection is based on the weight computed over V elements.

As for the MCP-sz and MCTTP-sz algorithms, we apply MCP and MCTTP, respectively, within each subzone. Being N/B (on average) the number of intersections within each subzone, the computational complexity of these algorithms results to be $O(B \log B \times V(\frac{N}{B})^k)$.

Finally, the complexity of the algorithm to solve the 0–1 KP is $O(VN + N \log N)$, since we just have to consider each of the N intersections and sort the values to obtain the best k choices.

5. 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.

Here, 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.

5.1. Scenario

For our performance evaluation, we selected real-world road topologies from the canton of Zurich, in Switzerland. Realistic traces of the vehicular mobility in such region are available from the Simulation and Modeling Group at ETH Zurich [3]. These traces describe the individual movement of cars through a queue-based model calibrated on real data [15]: they thus provide a realistic representation of vehicular mobility at both microscopic and macroscopic levels.

We considered the four road topologies depicted in Fig. 4, representing 100 km^2 portions of the urban areas centered at the cities of Zurich, Winterthur, Baden and Baar. For each topology, we extracted an hour and a half of vehicular mobility, in presence of average traffic density conditions. The number of road intersections in each scenario, and the amount of vehicles traveling within it during the observation time (after the filtering discussed below) are reported in Table 1.

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 1 min in the considered region. Fig. 3 shows where the filtering thresholds fall with respect to the cumulative distribution functions of visited intersections and trip duration, for each scenario. The selected thresholds result in a low percentage of cars being removed from the traces of the scenarios characterized by a higher traffic density (Zurich and Winterthur), while the filtering is heavier on the traces of the more rural scenarios (Baden and Baar), where the conditions set above are harder to meet. However, the resulting numbers, in Table 1, still guarantee the statistical validity of the tests conducted over all road topologies.

Table 1
Road topologies parameters.

	Zurich	Winterthur	Baden	Baar
Intersections	83	43	38	46
Vehicles	70,537	13,578	11,632	9876

5.2. Maximizing contacts

In the scenario described above, we first run the deployment algorithms for contact maximization presented in Section 4.1. The metric we are interested in is 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.

The selected settings for the MCP-sz algorithm were $L = 4$, $k_l = k$. This choice was the result of calibration tests run for MCP-sz on the different road scenarios, whose outcome is shown in Fig. 5. There, we can notice how the impact of both the value of L and of the expression of k_l on the coverage ratio is very small, for the values of k (i.e., number of DPs deployed) that we are interested in. We thus picked $L = 4$, since it implies a stronger locality of decisions and thus a reduced computational complexity, and $k_l = k$, as a simpler yet efficient choice with respect to more composite formulations.

Also, in order to provide a lower-bound benchmark to the performance of the schemes, we tested the performance of a random deployment, that ignores the vehicular mobility information and whose outcomes result from averaging multiple tests over each road topology.

The coverage ratio achieved by the different contact maximization schemes is shown in Fig. 6, for each street layout. For each deployment algorithm, the ratio is recorded versus the number of allowed DPs k.

Three different behaviors can be distinguished in all the scenarios considered. The first is that of the random algorithm, which, lacking all information on the movement of vehicles, performs poorly: it needs a large number of DPs (typically more than 50%

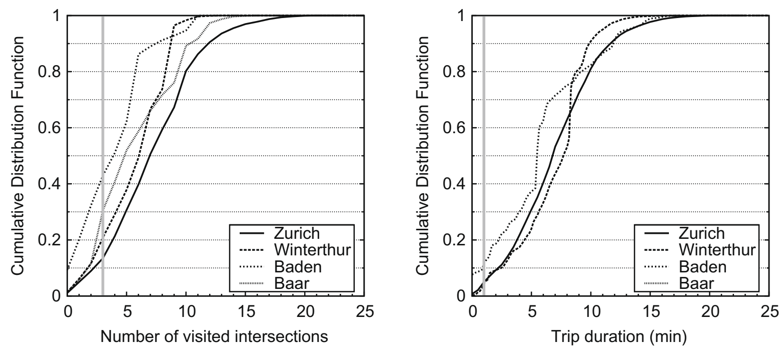


Fig. 3. CDF of the number of intersections traversed (left) and of the trip duration (right) for all vehicles in the four traces.

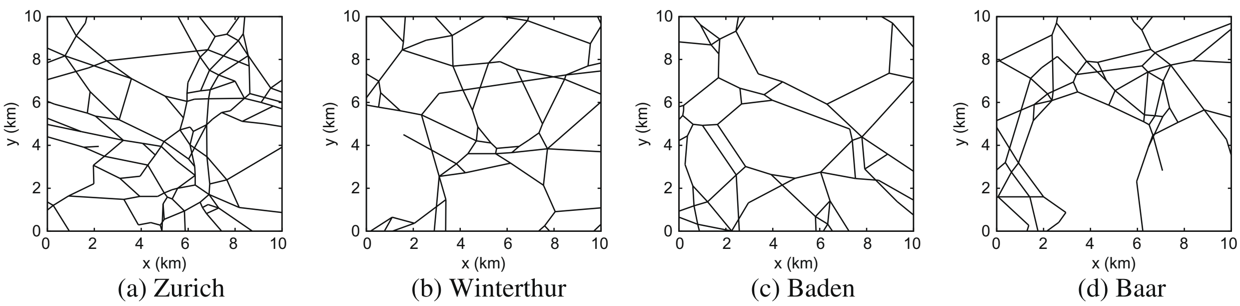
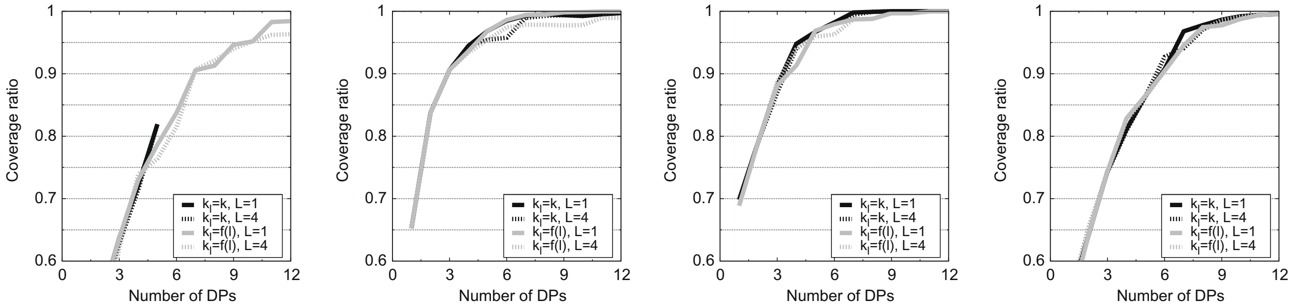
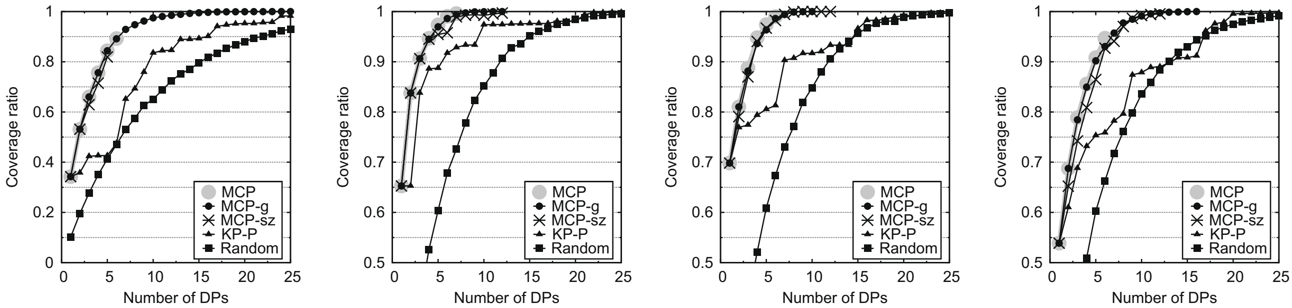


Fig. 4. Road topologies layouts.



(a) Zurich (b) Winterthur (c) Baden (d) Baar

Fig. 5. MCP-sz calibration. In the label of each plot, $k_l = f(l)$ stands for $k_l = \lceil \frac{k}{2^{L-l}} \rceil + 2^{L-l} - 1$ (see Section 4.1.2).



(a) Zurich (b) Winterthur (c) Baden (d) Baar

Fig. 6. Ratio of vehicles experiencing at least one contact with a DP versus the number k of DPs deployed, for each road topology.

of intersections) to be deployed in order to provide one contact or more to each vehicle.

The second behavior is that of the KP-P 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-P algorithm performs better than the random one, although its absolute result still has wide margins of improvement. As a matter of fact, its CDFs grow faster than the random one, but do not reach again a coverage ratio equal to one until almost half of intersections are covered by one DP. Moreover, the progress in terms of covered vehicles is quite irregular as the number k of deployed DPs grows. This suggests that the KP-P scheme can at times deploy DPs at new intersections 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-P algorithm is extremely high, since the greedy and subzone schemes cover 90% of vehicles with DPs covering between 5% and 10% of the available intersections, and 100% of vehicles with DPs at around 15% of intersections.

The variability in the percentages above is due to the different scenarios we consider. We also note that, although the relative performance of the algorithms is unaffected by the different street layouts, the absolute coverage ratios change (note the different ranges in the y axes of the plots). More precisely, the higher complexity of the road topology in the Zurich area results in lower coverage ratios when just a few DPs are present, whereas a single DP is sufficient to already cover more than 50% of cars in the other three scenarios. Indeed, in a smaller urban center, most traffic tends to gather over one or two main roadways, and it is thus easier to cover by placing DPs at strategic intersections. In larger metropolitan areas, instead, the vehicular flows split over a number of major routes, which makes it harder to cover them with a limited number of DPs.

Finally, we stress that the brute force solution to the MCP problem is computationally feasible only for low values of k, for which a 100% coverage cannot be usually reached. On the other hand, heuristics can be run even when DPs are deployed at a high percentage of the available intersections. Arguably, the fact that MCP-g and MCP-sz achieve a performance similar to that of MCP turns out to be an extremely important result, since it yields computationally-feasible quasi-optimal placements of DPs when a coverage close or equal to 100% is the goal of the deployment.

Further insight into the different behaviors is provided by Fig. 7. The figure shows the actual positions of the DPs over the Zurich 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-P 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.

By summarizing the results, we can conclude that:

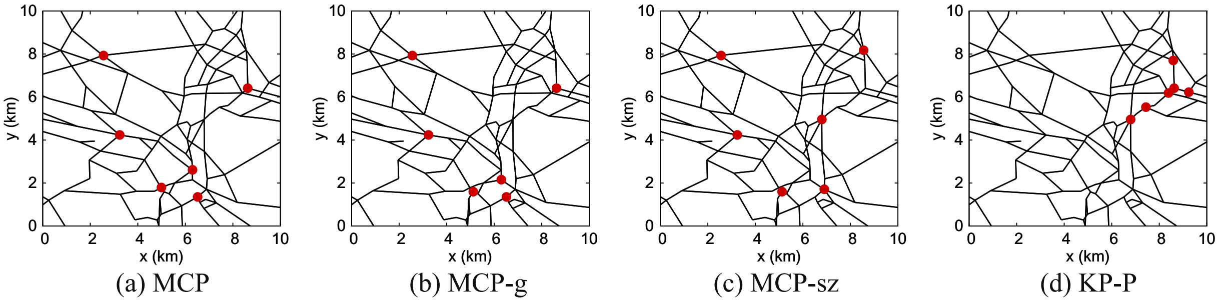


Fig. 7. Zurich road topology: deployments of DPs obtained with different algorithms maximizing contacts, for $k = 6$.

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

5.3. 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. Thus, in this case we are not only interested in the coverage ratio as a metric, but also in the *coverage time*, i.e., the amount of time that each vehicle spends within range of DPs during its trip in the considered scenario. Once more, a random deployment is employed to benchmark the performance of the algorithms we introduced in Section 4.2.

The coverage ratio achieved by such algorithms, in different road topologies and as the number of deployed DPs k varies, is depicted in the plots of Fig. 8, for a time threshold value $\tau = 30$ s. Exactly as observed in the previous section, also in this case the 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 the vehicles. The MCTTP, MCTTP-g, and MCTTP-sz solutions, leveraging their knowledge of cars trajectories, guarantee the highest coverage and tend to perform similarly. Such result is consistent through all scenarios, although the entity of the difference in the coverage ratio provided by the diverse deployment algorithms varies with the road topology considered: also in this case, a more complex road topology, such as that of Zurich, leads to more significant differences between the schemes that are mobility-aware and those that are not.

Fig. 9 reports instead the distribution of the coverage time, for $\tau = 30$ s (as remarked by the vertical threshold line in the plots), and in the specific case in which $k = 6$ DPs are deployed over each road topology. The goal now is to maximize the time spent by vehicles under coverage of DPs, up to the threshold τ seconds. The common result in all road topologies is that random deployments lead to small coverage times, whereas the other schemes tend to behave similarly, although KP-T is characterized by a more skewed distribution than those of MCTTP, MCTTP-g, and MCTTP-sz. As a matter of fact, the deployments determined by KP-T at a time result in more vehicles with very low coverage times, and more vehicles with very high coverage times. Conversely, MCTTP, MCTTP-g, and MCTTP-sz lead to more balanced distributions, where many vehicles experience a coverage time around the threshold τ. Once more, these observations hold for all the scenarios considered.

When comparing the coverage times in Fig. 9 with the coverage ratios in Fig. 8, we can notice that MCTTP, MCTTP-g, and MCTTP-sz provide very similar performance, which is superior to those achieved by the other schemes, for the considered settings $\tau = 30$ s and $k = 6$. Indeed, a random deployment of DPs induces both a lower number of vehicle-to-DP contacts and a shorter coverage time with respect to the solutions above. The KP-T solution leads to a performance comparable to those of MCTTP and relative heuristics in terms of coverage time, although with the skewness discussed before; however, this result is paid at a high coverage ratio cost.

When the value of the time threshold τ is increased, the constraint on the coverage time becomes stricter. Figs. 10 and 11 show, respectively, the coverage ratio (for varying k) and the CDF of the coverage time (for $k = 6$), when τ is set to 60 s. In the plot of Fig. 11, the threshold value τ is highlighted by the vertical dashed line. We can note how such coverage time threshold is very hard to reach with the limited number of DPs we bind our deploy-

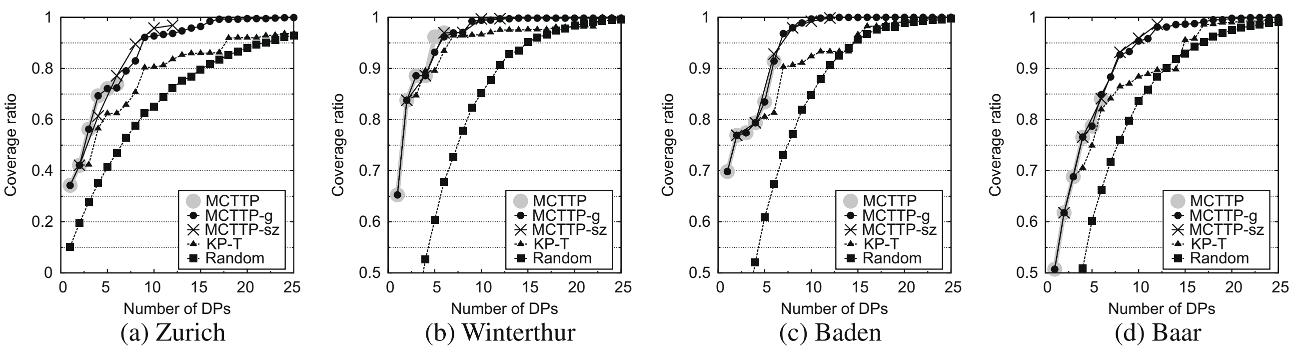
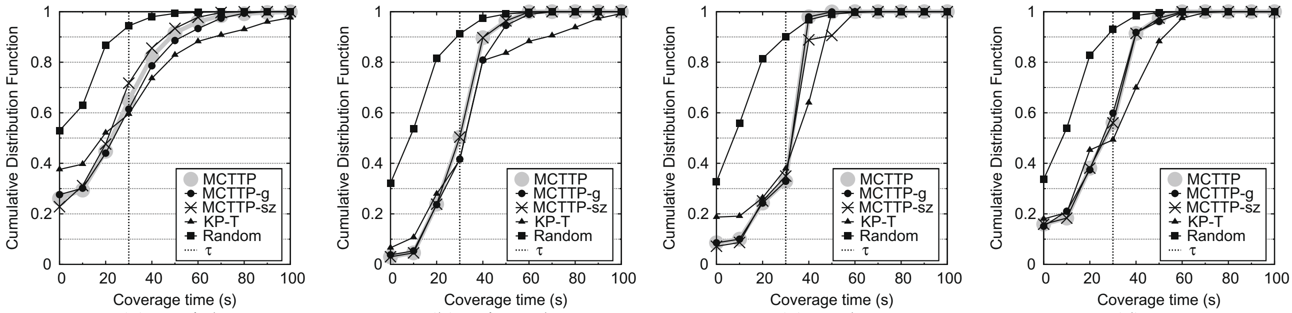


Fig. 8. Coverage ratio versus the number k of DPs deployed, for $\tau = 30$ s.



(a) Zurich (b) Winterthur (c) Baden (d) Baar

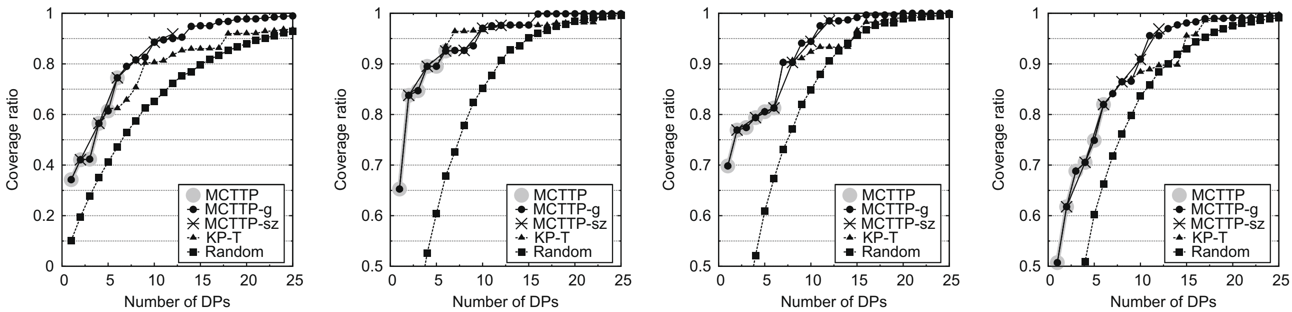
Fig. 9. Cumulative distribution function of the coverage time, for $\tau = 30$ s and k=6.

ment to: the percentage of vehicles covered for τ seconds when $k = 6$ is quite small, forcing the deployment of more DPs, or the increase of their dissemination range.

In any case, the random and KP-T schemes still perform worse than MCTTP and its greedy and subzone versions. In turn, the latter algorithms result in similar, but slightly reduced coverage ratios with respect to the case where smaller values of τ are considered, in an attempt to cover each vehicle for a longer time and match the threshold requirement.

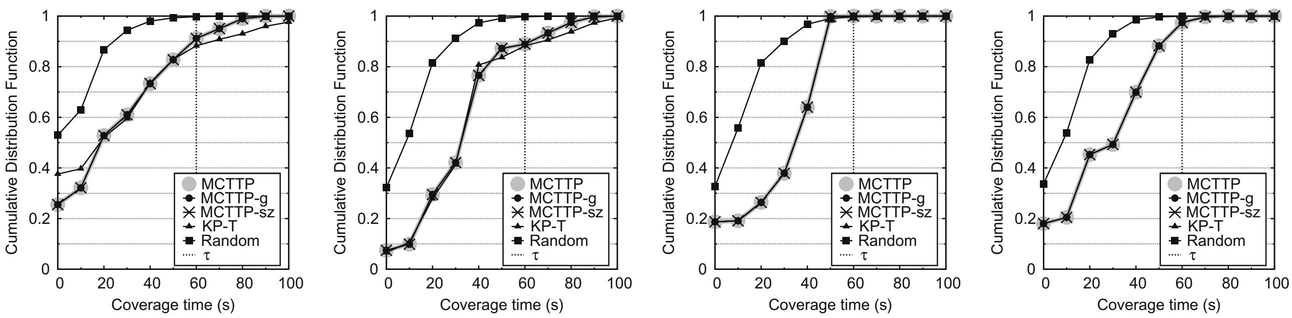
The relationship of the τ-dependent schemes from τ is studied in Fig. 12, for the complex Zurich road topology. There, we focus on MCTTP-g, since the other algorithms showed similar behaviors, and evaluate it as τ ranges between 5 and 120 s. The coverage ratio, in the left plot of Fig. 12, shows how 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 τ, 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. Instead, when τ 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 CDFs of the coverage time, on the right plot of Fig. 12, where the same behavior of the MCTTP-g algorithm is observed, as τ varies. It can be noted, however, how MCTTP-g with $\tau = 5$ s matches MCP-g in terms of coverage ratio, but outperforms it in terms of coverage time. Similarly, MCTTP-g with $\tau = 120$ s 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



(a) Zurich (b) Winterthur (c) Baden (d) Baar

Fig. 10. Coverage ratio versus the number k of DPs deployed, for $\tau = 60$ s.



(a) Zurich (b) Winterthur (c) Baden (d) Baar

Fig. 11. Cumulative distribution function of the coverage time, for $\tau = 60$ s and $k = 6$.

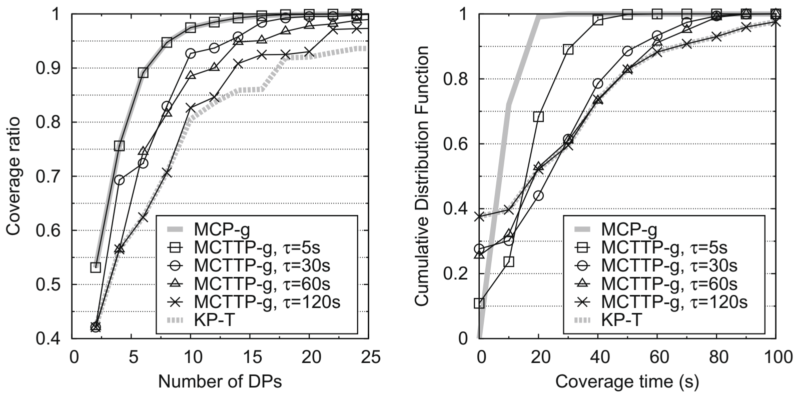


Fig. 12. Zurich road topology: coverage ratio versus the number k of DPs (left) and coverage time CDF, for $k = 6$ (right).

than contacts-only or time-only driven solutions even in borderline conditions.

We can draw the following conclusions:

- when the goal is to maximize 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 τ can be used to calibrate the deployment so that it is preferably driven by vehicle-to-DP contacts or by coverage time.

6. Related work

Wireless Access Point or Base Station placement is a well-known research topic, however most of the works that have addressed this problem so far consider a continuous infrastructure radio coverage. Wright [16] proposes a variant of the Nelder-Mead "simplex" method for finding optimal base station placement. Whitaker et al. [5] investigate macro-cell network planning in cellular networks and analyze the effects of cell density on the infrastructure cost of the network and the effects of increasing infrastructure expenditure on service coverage. Other similar works use set covering problems for network planning: Tutschku [17] introduces the Set Cover Base Station Positioning Algorithm that is based on a greedy heuristic for solving the Maximal Coverage Location Problem in radio networks, while Amaldi et al. [18] study WLAN positioning taking into account IEEE802.11 access mechanism. None of the former works however consider wireless networks with intermittent coverage, where base stations are planned taking into account non-continuous coverage for mobile users. Recently, Lochert et al. [19] have tackled the problem of sparse Road Side Units placement, formulating it as an optimization problem solvable with genetic algorithms. However, the goal of the deployment is in that case the aggregation of data on vehicular traffic conditions, and not the dissemination of information: the diverse target leads to a different problem formulation and, thus, solution. An interesting work is also presented in [20], where Chaintreau et al. focus on opportunistic communications between nodes that move between classes; classes may represent locations or states. The authors show that gossiping through opportunistic contacts (e.g., between taxi cabs) lead to efficient information update, and exploit their findings for base station deployment.

Our architecture assumes the InfoStation model [21], where small islands of coverage provide low-cost information services for mobile users. In the area of InfoStation models, the objective of most of the works that have appeared in the literature has been routing and reliability of information delivery, but not optimal AP placement. As an example, Sollazzo et al. [22] propose TACO-DTN, a content dissemination system that uses the InfoStation model, and study the routing governance and management of the AP resources when the AP disseminates information. Cohen et al. [23] consider the case where a set of Information Dissemination Devices disseminate information to passing mobile nodes. The authors use the Knapsack formulation to decide which messages should be broadcast by every dissemination device.

In the context of sensor networks, several studies [6–8] have considered the problem of deploying multiple sinks in presence of stationary nodes. The goal of these works, however, is to minimize energy consumption, while guaranteeing that sensors can access at least one sink through either single- or multi-hop communications. In the case of mobile sensors [10], the problem still differs significantly form ours, not only because of the different objectives, but also because of the different node mobility. Sensors mobility is typically represented by random models or follow predefined trajectories; also it is characterized by a much lower speed than vehicular mobility; it follows that, unlike our case, node movement can be predicted and exploited for sink deployment.

Finally, Maeda et al. [24] have proposed a technique to derive the movement of pedestrians in an urban scenario, starting from the street layout and observations of density at intersections. A similar technique could be exploited to derive (otherwise unavailable) information on the exact mobility of cars from measurements of congestion at crossroads, since we proved how knowledge of the former is much more useful than knowledge of the latter in order to achieve an optimal DPs deployment for information dissemination.

To summarize, our study differs from previous work in that it deals with optimal deployment of dissemination points when full coverage of the network area is not required and nodes have intermittent connectivity with the network infrastructure. Also, it addresses the case of generic user vehicles and, as is conceivable, it assumes that limited information is available on vehicle mobility as well as on their communication capabilities.

7. Conclusions

We proposed a maximum coverage approach to the problem of information dissemination in intelligent transportation systems. 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.

Acknowledgment

This work has been supported by the EuroNF NoE (FP7), through the Special Joint Research Project "InfoMob".

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, in: MobiCom, Montreal, Canada, 2007.
[2] J. Burgess, B. Gallagher, D. Jensen, B.N. Levine, Maxprop: routing for vehicle-based disruption-tolerant networks 2006, in: IEEE INFOCOM, Barcelona, Spain.
[3] ETH traces, <http://lst.inf.ethz.ch/ad-hoc/car-traces>.
[4] S.O. Krumke, M.V. Marathe, D. Poensgen, S.S. Ravi, H.-C. Wirth, Budgeted maximum graph coverage, Lecture Notes In Computer Science 2573 (2002) 321–332.
[5] R.M. Whitaker, L. Raisanen, S. Hurley, The infrastructure efficiency of cellular wireless networks, Computer Networks 48 (6) (2005) 941–959.
[6] H. Kim, Y. Seok, N. Choi, Y. Choi, T. Kwon, Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks, in: International Conference on Information Networking (ICOIN), 2005.
[7] E.I. Oyman, C. Ersoy, Multiple sink network design problem in large scale wireless networks, in: IEEE ICC, 2004.
[8] W.Y. Poe, J.B. Schmitt, Minimizing the Maximum Delay in Wireless Sensor Networks by Intelligent Sink Placement, Technical Report 362/07, University of Kaiserslautern, Germany, 2007.
[9] L. Qiu, R. Chandra, K. Jain, M. Mahdian, Optimizing the placement of integration points in multi-hop wireless networks, in: International Conference on Network Protocols (ICNP), Berlin, Germany, 2004.
[10] Y. Hu, Y. Xue, Q. Li, F. Liu, G.Y. Keung, B. Li, The sink node placement and performance implication in mobile sensor networks, Journal on Mobile Networks and Applications (MONET) 14 (2009) 230–240.
[11] VanetMobiSim, <ttp://vanet.eurecom.fr>.
[12] M. Fiore, J. Härri, F. Filali, C. Bonnet, Vehicular mobility simulation for vanets, in: IEEE/SCS Annual Simulation Symposium, 2006.
[13] A.A. Ageev, M.I. Sviridenko, Approximation algorithms for maximum coverage and max cut with given sizes of parts, Lecture Notes in Computer Science 16 (1999).
[14] D. Pisinger, Where are the hard knapsack problems?, Computers and Operations Research 32 (2004) 2271–2284
[15] N. Cetin, A. Burri, K. Nagel, A Large-scale Multi-agent Traffic Microsimulation Based on Queue Model, STRC'03, Ascona, Switzerland, 2003.
[16] M.H. Wright, Optimization methods for base station placement in wireless applications, in: IEEE VTC, Ottawa, Canada, 1998.
[17] K. Tutschku, Demand-based radio network planning of cellular mobile communication systems, in: IEEE INFOCOM, San Francisco, USA, 1998.
[18] E. Amaldi, A. Capone, M. Cesana, F. Malucelli, Optimizing WLAN radio coverage, in: IEEE ICC, Madison, USA.
[19] C. Lochert, B. Scheuermann, C. Wewetzer, A. Luebke, M. Mauve, Data Aggregation and Roadside Unit Placement for a Vanet Traffic Information System, ACM VANET, S. Francisco, USA, 2008.
[20] A. Chaintreau, J.Y. Le Boudec, N. Ristanovic, The Age of Gossip: Spatial Mean-Field Regime, ACM SIGMETRICS 2009, Seattle, WA, 2009.
[21] R.H. Frenkiel, B.R. Badrinath, J. Borras, R. Yates, The infostations challenge: balancing cost and ubiquity in delivering wireless data, IEEE Personal Communications Magazine 7 (2) (2000) 66–71.
[22] G. Sollazo, M. Musolesi, C. Mascolo, TACO-DTN: a time-aware content-based dissemination system for delay tolerant networks, in: ACM MobiOpp, Puerto Rico, USA, 2007.
[23] R. Cohen, D. Raz, M. Aezladen, Locally vs. globally optimized flow-based content distribution to mobile nodes, in: IEEE INFOCOM, Rio de Janeiro, Brazil, 2009.
[24] K. Maeda, A. Uchiyama, T. Umedu, H. Yamaguchi, K. Yasumoto, T. Higashino, Urban pedestrian mobility for mobile wireless network simulation, Ad Hoc Networks 7 (2009) 153–170.