

Information Dissemination in VANETs: Deployment Strategies for Maximizing Coverage

O. Trullols, J.M. Barcelo-Ordinas
Universitat Politècnica de Catalunya
Barcelona, Spain
Email: <trullols.joseb>@ac.upc.edu

M. Fiore, C. Casetti, C.-F. Chiasserini
Politecnico di Torino
Torino, Italy
Email: <marco.fiore,casetti,chiasserini>@polito.it

Abstract—We consider that a given number of Dissemination Points (DPs) have to be deployed for disseminating information to vehicles travelling in an urban area. We 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. 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.

Keywords—Vehicular networks; network planning; maximum coverage

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 vehicles-to-infrastructure (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. 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). The problem is NP-hard, thus we propose heuristic algorithms to solve it 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 an 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 the number of vehicles, among the possible V , served (i.e., covered) by the 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.

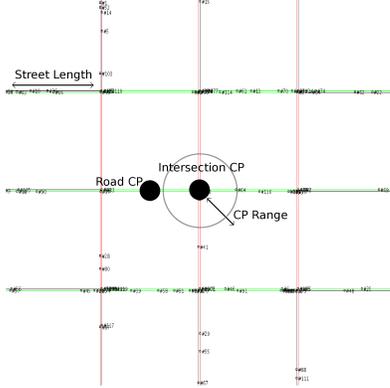


Figure 1. road topology

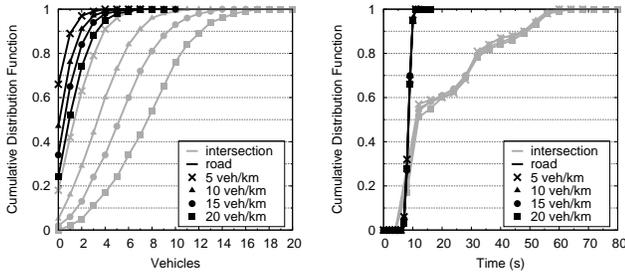


Figure 2. CDF of the number of vehicles within range of the DP (left) and of the time spent by vehicles within range of the DP (right), with $R = 50$ m and for different vehicle densities

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, shown in Fig. 1, 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 vehicles movement is simulated with Vanet-MobiSim [5], using 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].

We considered different vehicular lane densities, ranging between 5 and 20 vehicles/km, which map onto 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 DP's dissemination range R : a higher number of vehicles, as well as longer dwelling times, correspond to a

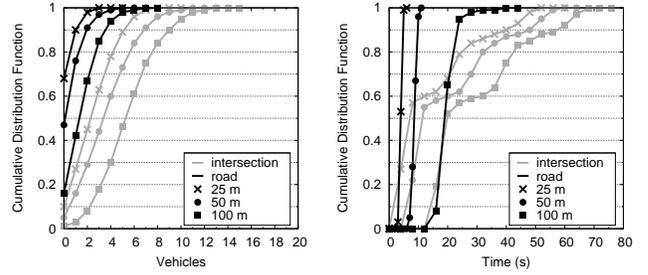


Figure 3. CDF of the number of vehicles within range of the DP (left) and of the time spent by vehicles within range of the DP (right), with a vehicular lane density of 10 vehicles/km and for different DP dissemination ranges

higher potential for information dissemination, as more users can receive larger portions of the content provided by the DP. Fig. 2 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. 3 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 reminder of the paper, we will focus on the problem of DPs deployment at intersections of the road topology.

IV. DEPLOYMENT ALGORITHMS FOR 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

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

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-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 $\mathcal{S} = \{S_1, S_2, \dots, S_N\}$, where 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, which requires global knowledge of the network system, and then we propose a localized, hierarchical heuristic called *the subzone algorithm*.

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. 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

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 $U = \emptyset$
-

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 [9]. 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 ν_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 name this algorithm 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).

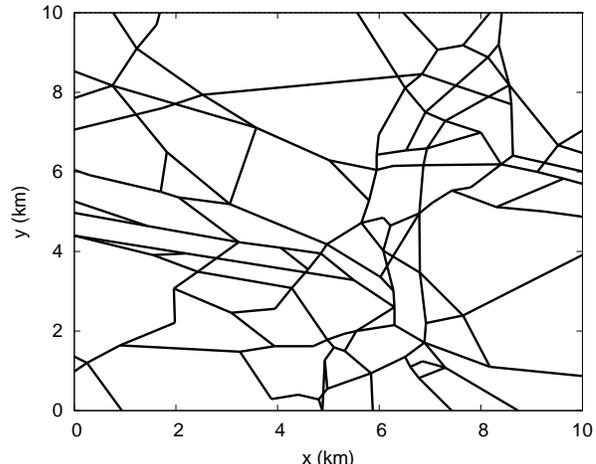


Figure 4. Performance evaluation road topology

V. PERFORMANCE EVALUATION

We applied the algorithms presented in the previous section 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.

A. Scenario

For our performance evaluation, we set the dissemination range of every DP to 100 m and 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 [11]. These traces describe the individual movement of cars through a queue-based model calibrated on real data [12]: they thus provide a realistic representation of vehicular mobility at both microscopic and macroscopic levels. The road topology we considered is a 100 km² portion of the urban and suburban area of the city of Zurich, a map of which can be found in the plots of Fig. 4. We extracted an hour and a half of vehicular mobility, in presence of average 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. 5 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.

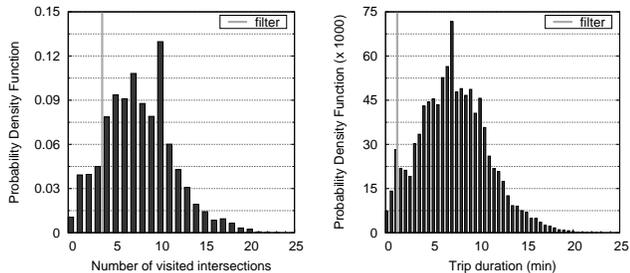


Figure 5. PDF of the number of intersections traversed (left) and of the trip duration (right) for all vehicles in the trace

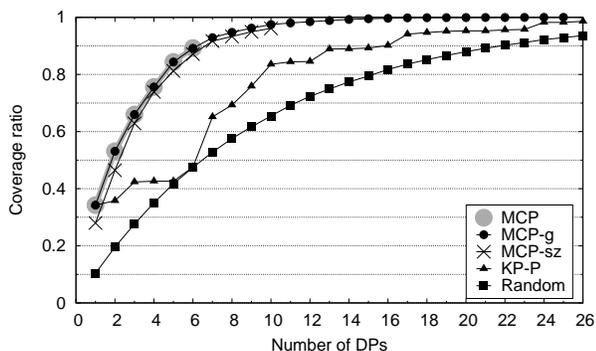


Figure 6. Ratio of vehicles experiencing at least one contact with a DP versus the number k of DPs deployed, for different algorithms that maximize contacts

B. Maximizing contacts

In the scenario described above, we run the deployment algorithms for contact maximization. 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. 6. 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 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 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. 7. The figure shows the actual positions of the DPs over the road topology¹, 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. 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 km²;
- 3) exploiting these properties it is possible to inform a high percentage of vehicles by deploying DPs at a small percentage of intersections.

VI. CONCLUSION

We proposed a maximum coverage approach to the problem of information dissemination in vehicular networks.

¹The light blue dots are used to mark the DPs' position, their size is not related to the DP's dissemination range.

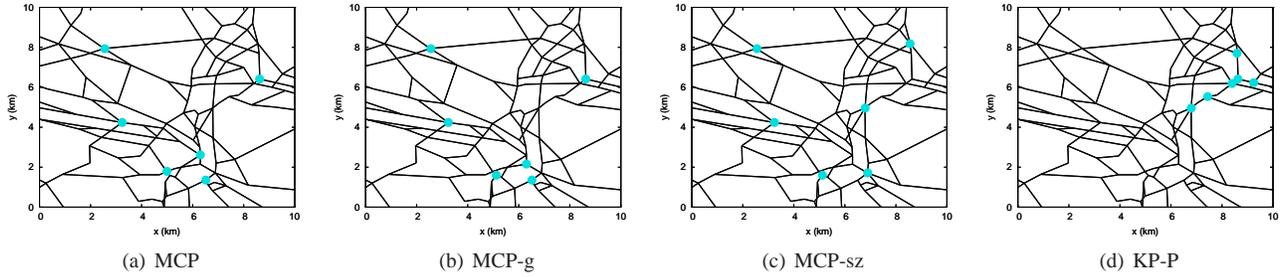


Figure 7. Deployments of DPs obtained with different algorithms maximizing contacts over the Zurich road topology, for $k = 6$

The formulations and relative heuristics we presented tackle the case in which maximizing vehicle-to-DP contacts is our goal. We evaluated the different maximizing contacts's 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.

VII. FUTURE WORK

In this paper we have assumed that the information is a small and self-contained item, which vehicles obtain just contacting a DP once. As a future work, we want to find new strategies that in addition to deploy the DPs maximizing the nodes covered, also 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. New deployment strategies should be developed considering the case in which vehicle-to-DP contact times have an impact on the dissemination process. In this case the problem gets a different formulation, which aims at favouring both the number of contacted vehicles as well as the time for which they are covered.

ACKNOWLEDGMENT

This work was partially supported by the EU through the EuroNF NoE, by Spanish Ministry of Science and Technology under grant TSI2007-66869-C02-01 and partially by Regione Piemonte through the VICSUM project.

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. Härri, 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] G. E. Whitehouse, G. W. Depuy, R. J. Moraga, "Meta-RaPS Approach for Solving the Resource Allocation Problem," *Biannual World Automation Congress*, pp. 103–108, 2002.
- [9] D. Pisinger, "Where Are the Hard Knapsack Problems?," *Technical Report 2003/08*, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark, 2003.
- [10] S. Khuller, A. Moss, J. Naor, "The Budgeted Maximum Coverage Problem," *Information Processing Letters*, vol. 70, no. 1, pp. 39–45, 1999.
- [11] ETH traces, <http://lst.inf.ethz.ch/ad-hoc/car-traces>.
- [12] N. Cetin, A. Burri, K. Nagel, "A Large-scale Multi-agent Traffic Microsimulation Based on Queue Model," *STRC'03*, Ascona, Switzerland, 2003.