Non-Intrusive Planning the Roadside Infrastructure for Vehicular Networks

Cristiano M. Silva, *Member, IEEE*, Wagner Meira, Jr., *Member, IEEE*, and João F. M. Sarubbi, *Member, IEEE*

***Abstract*—In this article, we describe a strategy for planning the roadside infrastructure for vehicular networks based on the global behavior of drivers. Instead of relying on the trajectories of all vehicles, our proposal relies on the migration ratios of vehicles between urban regions in order to infer the better locations for deploying the roadside units. By relying on the global behavior of drivers, our strategy does not incur in privacy concerns. Given a set of α available roadside units, our goal is to select those α-better locations for placing the roadside units in order to maximize the number of distinct vehicles experiencing at least one V2I contact opportunity. Our results demonstrate that full knowledge of the vehicle trajectories are not mandatory for achieving a close-to-optimal deployment performance when we intend to maximize the number of distinct vehicles experiencing (at least one) V2I contact opportunities.**

***Index Terms*—Access infrastructure, VANET, roadside units, deployment, infrastructure design, V2I.**

I. Introduction

INTELLIGENT Transportation Systems are grounded on vehicular networks [1], i.e., sophisticated communication networks receiving data from several entities composing the traffic system. In a vehicular network, the communication may happen in an *ad hoc* basis where vehicles exchange messages without any support infrastructure [2].

However, the *ad hoc* communication may become inefficient in sparse areas such as highways, rural zones, and low peak hours in the city due to the lack of communicating pairs and radio obstacles. The intense mobility of vehicles also makes routing far complicated as we lack reliable means to infer the future location of vehicles. Although the communication may take place in an *ad hoc* basis, several works [3], [4] demonstrate that a minimum support infrastructure may largely improve the overall efficiency of the network.

Additionally, vehicles do not necessarily demand continuous coverage since they may engage in "infofueling" as they opportunistically drive past the roadside units [5]. Thus, when planning the roadside infrastructure for vehicular networks we must prioritize a subset of locations for receiving the infrastructure. When we review the previous works we notice the existence of two clusters of deployment strategies. The first cluster holds strategies very similar to the deployment of base stations in cellular networks (deployment at the densest locations) without taking the mobility into account.

The second cluster holds the modern deployment strategies based on previous-and-full knowledge of the trajectories of all vehicles using realistic car mobility traces. However, when considering a real deployment, the assumption of full knowledge of the trajectories of all vehicles seems unrealistic since it raises several privacy concerns. Furthermore, processing the trajectories of all vehicles requires a very large computing effort, and gathering such information is far from trivial.

Differently from the previous approaches, we propose a mobility-driven strategy for the deployment of infrastructure for vehicular networks based on the global behavior of drivers. Instead of relying on the individual trajectories of vehicles, we divide the urban area into a set of adjacent urban cells and we use i) the migration ratios of vehicles between distinct urban cells, and ii) the density of vehicles within every urban cell to locate the roadside units. Among several optimization targets, we maximize the number of distinct vehicles contacting the infrastructure, an interesting metric when we intend to collect and disseminate small traffic announcements [6].

The **Full Projection of the Flow** (FPF) strategy is a refinement of our strategy proposed in [7]. The Full Projection of the Flow has the goal to estimate the expected number of vehicles not presenting V2I contact opportunities, given a set of deployed roadside units. By knowing the number of vehicles inside each urban cell and the migration ratios between every pair of urban cells we are able to select the most promising urban cells for receiving roadside units.

We evaluate our deployment algorithm considering realistic conditions (real road network and realistic flow), and we compare FPF to the greedy solution of the Maximum Coverage Problem (MCP-g), a close-to-optimal deployment strategy demanding previous knowledge of the trajectories of all vehicles for defining the location of the roadside units. Considering the realistic scenario, MCP-g provides contact opportunities for 91.9% of all vehicles, while the optimal deployment solved using an Integer Linear Programming Formulation on IBM CPLEX Optimizer[1] achieves 93.3%, a difference of just 1.4%.

The results demonstrate that previous knowledge of the trajectories of the vehicles is not required for achieving a close-to-optimal deployment performance: by covering 1.0% of the Cologne's road network, our FPF deployment yields

Manuscript received January 28, 2015; revised August 20, 2015; accepted September 27, 2015. Date of publication November 5, 2015; date of current version March 25, 2016. The Associate Editor for this paper was L. Li.

C. M. Silva is with the Departamento de Tecnologia, Universidade Federal de São João del Rei, 36307-352 São João del Rei-MG, Brazil (e-mail: cristiano@ufsj.edu.br).

W. Meira Jr. is with the Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte-MG, Brazil (e-mail: meira@dcc.ufmg.br).

J. F. M. Sarubbi is with the Departamento de Computação, Centro Federal de Educação Tecnológica de Minas Gerais, 35503-822 Divinópolis-MG, Brazil (e-mail: joao@decom.cefetmg.br).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2015.2490143

[1] http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

89.8% of all vehicles experiencing at least one V2I contact opportunity. On the other hand, if we assume previous knowledge of the individual vehicles trajectories, we improve FPF in just 2.3%.

This work is organized as follows: Section II presents the related work. Section III presents a strategy for representing road networks. Section IV presents the Full Projection of the Flow strategy. Section V presents the baselines: we compare our approach to MCP-g (deployment strategy relying on the full knowledge of the trajectories of the vehicles), and MCP-kp (deployment strategy that allocates the roadside units at the densest locations of the road network). Section VI presents the experiments considering the realistic scenario of Cologne, Germany. Section VII concludes the work.

II. Related Work

The deployment of infrastructure for vehicular networks has gained a lot of attention over the past years. Some researchers have devoted their attention in elaborating analytical formulations. The work [8] proposes the definition of an infrastructure for vehicular networks based on the conventional definition of transport capacity. The authors develop a mathematical model where the destination nodes are chosen at random by the source nodes, and they study the effect of the infrastructure node deployment in the capacity of vehicular networks using analytical expressions.

Alpha Coverage [9] provides worst-case guarantees on the interconnection gap while using fewer roadside units. A deployment of roadside units is considered α-covered if any simple path of length α on the road network meets at least one roadside unit. The authors compare the α-coverage with a random deployment of roadside units.

Other authors have focused on elaborating deployment strategies for several scenarios and objective functions. In [10] the authors propose a placement strategy for a set of access points to access a single gateway in an open space, similar to the base station placement in cellular systems, and the goal is to minimize the power consumption and the average number of hops from access points to gateways under the assumption of full coverage by roadside units.

In [11] the authors present the evaluation of a deployment strategy through the contact opportunity. The contact opportunity measures the fraction of distance or time that a vehicle is in contact with the infrastructure. The authors argue that such a metric is closely related to the quality of the data service that a mobile user might experience while driving.

In [12] the authors propose a greedy heuristic to place the roadside units aiming to improve vehicles connectivity while reducing the disconnections. The heuristic counts the amount of reached vehicles by each intersection considering the transmission range of the roadside units. Each intersection is considered as a potential roadside unit location. The optimal locations are selected based on the number of vehicle reports (per minute locations reported by taxis) received within the communication range of each roadside unit.

In [13] the authors propose three optimal algorithms for allocating the roadside units: Greedy, Dynamic and Hybrid algorithms. They assume: i) placing roadside units preferentially at important intersections; ii) allocating roadside units until every intersection is covered; and iii) distributing roadside units as even as possible. The importance of each intersection is evaluated using traffic factors including vehicle density, intersection popularity, and intersection particularity. The greedy algorithm simply deploys the roadside units at intersections in descending order of the intersection priority. The dynamic algorithm concentrates on the even distribution of the roadside units, while the hybrid algorithm combines both the greedy and the dynamic methods.

In [14] the authors study the deployment of the roadside infrastructure in 2-D by formulating an optimization problem and solving it using integer linear programming. The proposed optimization framework takes into account the effect of buildings on signal propagation, LAN lines and road topology. In addition, Roadgate [15] addresses the placement of roadside units guaranteeing a probability of contact between vehicles and the infrastructure.

In terms of our specific works, in [7] we present the concept of a deployment based on the Probabilistic Maximum Coverage Problem. We use a more simplistic model where we keep track of the migration ratios only between adjacent urban cells. This work differs from [7] in the sense that we consider a more complete model for describing the mobility based on the migration ratios between all pairs of urban cells, and not only on adjacent urban cells.

Additionally, in [16] we propose the Delta Deployment, a QoS-based deployment. The QoS is specified in terms of the contact time and the number of vehicles experiencing such V2I contact time. The Delta Deployment is a specialization of the Delta Network proposed in [17].

Along this work we propose partitioning the road network into a set of same size urban cells. The partition of road networks has already been proposed in the literature. However, the authors employ the partitioning as an intermediate step to accomplish: (i) simplified mathematical modeling of the problem [18]; (ii) subdivision of the problem into several smaller ones [19]; or even (iii) simplified road network consisting of only horizontal and vertical roads [20]–[22].

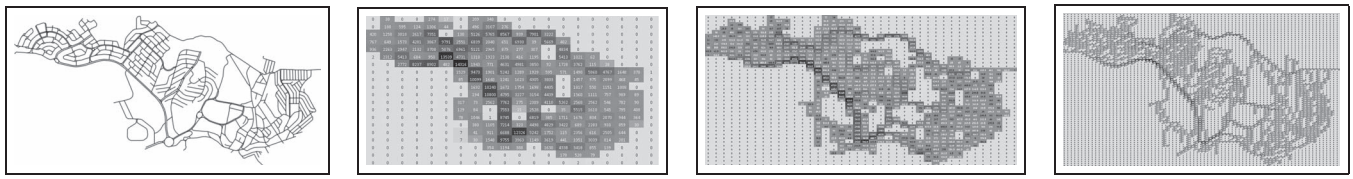
Our work differs from all the previous ones in the sense that we demonstrate that we do not need to track individual vehicles in order to achieve a close-to-optimal deployment performance when we intend to maximize the number of distinct vehicles presenting at least one V2I contact opportunity during the trip.

III. Representing Road Networks

We evaluate the deployment strategies considering a real road network. Before applying the deployment strategies, we must be able to represent road networks of arbitrary topology. In this study, we propose partitioning the urban area into a set of adjacent same size cells (i.e., grid model). Once the region is partitioned, we discard the original road network and we manipulate the flow between adjacent grid cells.

Partitioning allows us to represent the road network and its associated flow by a grid structure of arbitrary granularity. When we need more/less accuracy we simply increase/decrease the number of grid cells covering the region.

There are several motivations to partition a road network. Probably, the most important ones are: i) the reduction of the



(a) (b) (c) (d)

Fig. 1. Distinct grid setups. (a) Road network. (b) 20×20 grid. (c) 40×40 grid. (d) 80×80 grid.

computing efforts; ii) the flexibility on the resolution of the grid; iii) the abstraction of the flow.

We reduce the computing efforts because when we partition a road network we reduce the possible locations to be evaluated reducing the computing efforts.

Flexibility on the resolution of the grid allows us a trade-off between "resolution" and "processing time." A proper balance between such features should enable us for dynamic, and even, real-time deployment capabilities. Although dynamic deployment is not possible with stationary infrastructure, several works [23]–[25] propose the adoption of mobile roadside units to support the vehicular communication. Thus, dynamic deployment may guide the allocation of the mobile roadside units in order to address traffic fluctuations.

Abstraction of the flow is also a very interesting feature: a partitioned road network is implemented as a matrix of integers. Thus, the complexity for processing a large metropolitan area is the very same for processing a small town when both areas are partitioned using the same grid setup. The complexity is not related to the intensity of the flow, but to the selected grid setup.[2]

In Fig. 1(a) we present a real road network (Ouro Branco, Brazil). In Fig. 1(b)–(d) we show how such road network is modeled by grid setups ranging from 20×20 up to 80×80. Each urban cell indicates the number of vehicles crossing the cell during a given time interval. Each location is mapped to an urban cell and vehicles always move between adjacent urban cells. As we reduce the dimensions of the urban cells we increase the resolution of the grid making it increasingly similar to the original road network.

Fig. 1(b)–(d) indicates a rectangular shape for the urban cells, which may seem a little bit unusual. In fact, our goal is not to represent ranges of coverage but to divide-and-conquer: we divide the city into urban cells and then we select those α-cells resulting in a better coverage. The exact location of each roadside unit inside a given cell is out of our scope since we must take into account several practical issues (availability of energy supply, signal interference, large constructions blocking the signal, and so forth).

IV. FPF: Full Projection of the Flow Deployment

The FPF strategy uses the migration ratios in order to select the urban cells that must receive the roadside units. The goal is to identify the number of uncovered[3] vehicles along the entire road network. Initially, we select the urban cell presenting the highest number of vehicles for receiving a roadside unit. Then we project the flow, i.e., all the flows heading to the roadside unit and all the flows leaving the roadside unit are removed. Then, we perform another algorithm iteration (select the densest location followed by the projection of the flow) until we select α urban cells.

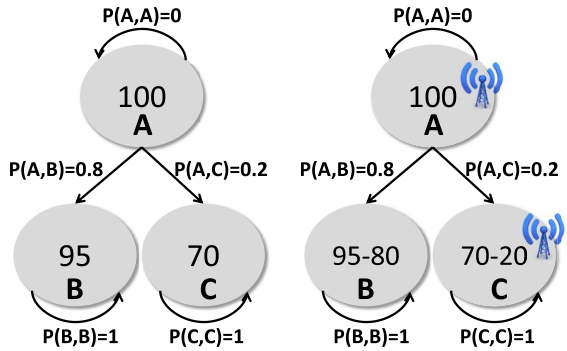


Fig. 2. Selection of urban cells for receiving the roadside units. (a) Scenario. (b) FPF strategy.

Now we illustrate FPF. Fig. 2(a) presents a road network partitioned into three locations {A, B, C}. Inside each location we indicate the number of vehicles crossing the location during the observation period. We also indicate the migration ratios. For instance, given that a vehicle is at A at time t, it has 80% probability of reaching B at time $t+1$. Thus, $\mathrm{P(A, B)} = 0.8$.

Our goal is to select two locations for receiving roadside units. FPF selects the locations {A, C}. Initially, FPF selects the location presenting the highest traffic {A}. Then, it projects the flow. From the 95 vehicles recorded in {B}, 80 vehicles are already covered in {A}. Thus, selecting {B} implies in covering just another new 15 vehicles. Similarly, from the 70 vehicles recorded in {C}, 20 vehicles arrive from {A}. Thus, deploying a roadside unit into {C} implies in covering another 50 uncovered vehicles, and FPF selects {A,C}.

The FPF strategy is presented in Algorithm 1: it receives as input the bi-dimensional matrix M of urban cells representing the flow within a partitioned road network, the number of available roadside units (α), the four dimensional matrix of migration ratios (P), and the grid setup (ϕ).

Since M has ϕ rows and ϕ columns, $|M| = \phi^2$. Each $M_{i,j}$ indicates the number of vehicles crossing the urban cell $\{i, j\}$ during some given time interval. Complementary, each $P_{i,j,k,l}$ indicates the migration ratios between the urban cells $\{i, j\}$ and $\{k, l\}$, considering the same time interval.

We assume both the density of vehicles and the migration ratios as accurate information since they represent the algorithm

[2]For instance, 20×20, 30×30, and so on.

[3]We are assuming a slightly different concept of coverage: traditional usage of coverage indicates a continuous region where users are supposed to meet connection. But, because we are assuming **infostations**, we consider small islands of coverage: fragmented and possibly disconnect areas where users are supposed to meet connection.

input. Although we have not characterized the effects of deviations on the migration ratios, they impact the performance of our strategy. A small deviation will present a small impact on the performance of FPF possibly changing the location of some roadside units to some neighbor urban cells. A high deviation compromises the FPF strategy.

For the purposes of this work, we obtain the migration ratios and the density of vehicles along the road network by inspecting the mobility trace. When considering a real deployment we may not have the migration ratios available. In order to overcome this situation we may perform an incremental deployment of roadside units in order to gain the density of vehicles per urban cell and the migration ratios. As we increase the available mobility information, we proceed on deploying more roadside units iteratively, according to FPF.

Algorithm 1 FPF: Full Projection of the Flow Deployment.

Input: M, α, P, ϕ;
Output: Γ (solution set);
1: $\Gamma \leftarrow \emptyset$;
2: m $\leftarrow$GETCELLPRESENTINGMAXVEHICLES(M);
3: $\Gamma \leftarrow \Gamma \cup$ m;
4: **for** $i = 2$ to α **do**
5: FULLFLOWPROJECTION(m.x,m.y);
6: m $\leftarrow$GETCELLPRESENTINGMAXPROJECTION(M);
7: $\Gamma \leftarrow \Gamma \cup$ m;
8: **end for**
9: **return** Γ;
10: **procedure** FULLFLOWPROJECTION(x,y)
11: **for** i $= 1$ to ϕ **do**
12: **for** j $= 1$ to ϕ **do**
13: $M_{i,j} \leftarrow$ MAX$(M_{i,j} \times (1 - P_{i,j,x,y}), 0)$;
14: $M_{i,j} \leftarrow$ MAX$(M_{i,j} - M_{x,y} \times P_{x,y,i,j}), 0)$;
15: **end for**
16: **end for**
17: **end procedure**

Algorithm 1 presents FPF: it iteratively selects the urban cells presenting the highest projection of the flow. In the first iteration, FPF selects the densest urban cell to receive a roadside unit. Such urban cell is added to the solution set Γ. When FPF selects an urban cell, it recomputes the number of distinct-and-uncovered vehicles for all the remaining urban cells. Basically, FPF removes all flows heading to the selected urban cell. Then, FPF removes all flows leaving the urban cell using the migration ratios P.

The function *FullFlowProjection(x, y)* receives the coordinates of the selected urban cell {x, y} and removes vehicles that will cross the new roadside unit. The line 13 remove vehicles crossing {i, j} and then {x, y}. The line 14 removes vehicles crossing {x, y} and then {i, j}. Analytically,[4] we may represent the projection of the flow using Equations (1) and (2): Equation (1) excludes from $M_{i,j}$ vehicles crossing {i, j} and then {x, y}, where x and y are the grid coordinates of the urban cell receiving the new roadside unit and i and j represent all the urban cells. The Equation (2) excludes from $M_{i,j}$ vehicles crossing {x, y} and then {i, j}.

$$M_{i,j} \leftarrow M_{i,j} \times (1 - P_{i,j,x,y}) \quad (1)$$

$$M_{i,j} \leftarrow M_{i,j} - M_{x,y} \times P_{x,y,i,j}. \quad (2)$$

Since every time we deploy a roadside unit we must update the number of uncovered vehicles along the entire grid, the computational complexity of FPF is $\Theta(\alpha.M)$.

When computing the projection of the flow of vehicles the influence of distant vehicles decays fast indicating that distant vehicles are less likely to reach the selected urban cell (they have several possible trajectories). On the other hand, the number of distant vehicles increases very fast as we move away from the selected urban cell. The relation among these magnitudes determines the urban cell receiving a roadside unit and defines the efficiency of the entire method.

A. Addressing Traffic Fluctuations

Although in this work we are focused on the deployment of stationary roadside units, here we present a few comments on employing FPF to address traffic fluctuations. Traffic fluctuates according to the type and time of day, weather conditions, road works, and accidents. Thus, an architecture composed just of stationary roadside units may not be able to properly handle all the flow variations.

Similarly, an architecture composed just of mobile roadside units may lack part of the robustness provided by the stationary infrastructure. Thus, hybrid architectures combining both stationary and virtual-and-mobile roadside units would be interesting solutions for vehicular networks.

During rush hours the major roads get congested and the drivers use secondary roads as an alternative for escaping the congestions. Thus, major roads are likely to receive stationary roadside units, since they are the most important roads.

On the other hand, secondary roads may receive temporary support from mobile-and-virtual roadside units, such as unmanned aerial vehicles launched by the stationary roadside units. In this situation we may use FPF to indicate the destination of the mobile roadside units.

Adapting FPF for guiding the deployment of mobile roadside units is somehow trivial. Besides the input parameters described in Algorithm 1, we must also consider: i) the set of deployed stationary roadside units; ii) the actual location of the mobile roadside units; iii) the number of mobile roadside units (α'); iv) the autonomy of each mobile roadside. Then, the adapted FPF iterates until finding α' locations for moving the mobile roadside units meeting the autonomy constraint.

B. Integer Linear Programming Formulation

The goal of FPF is to select those α urban cells presenting the highest number of uncovered vehicles. FPF may be expressed as an Integer Linear Programming Formulation (ILPF). Let K be the set of vehicles, where $K = \{1, 2, \ldots, k\}$. Let α denotes the number of available roadside units. Let C be the set representing all urban cells, where $C = \{1, 2, \ldots, c\}$. Let M be a binary variable representing the trajectories: $M_{c,k} = 1$ if the vehicle k crosses the urban cell c (0, otherwise).

[4]Theoretical analysis is presented in [26], pages 51–58.

Let A represents the urban cells receiving roadside units: $A_c = 1$ if the urban cell c receives a roadside unit (0, otherwise). Let V represents the covered vehicles: $V_k = 1$ when the vehicle k crosses any roadside unit (0, otherwise). Assuming these variables, we may model FPF as follows:

$$\max \sum_{n=1}^{k} V_n. \tag{3}$$

Subject to:

$$\sum_{\forall c \in C \mid M_{c,k}=1} A_c \leq \alpha \tag{4}$$

$$\sum_{\forall c \in C \mid M_{c,k}=1} A_c \geq V_k \qquad \forall k \in K \tag{5}$$

$$V_k \geq A_c \qquad \forall c \in C, \quad \forall k \in K | M_{c,k} = 1. \tag{6}$$

The objective function 3 maximizes the number of distinct vehicles reaching roadside units. The constraint 4 ensures that the number of selected urban cells is less than α. The constraint 5 ensures that, whenever a vehicle k is covered, then k has crossed at least one urban cell owning a roadside unit. Finally, the constraint 6 ensures that, whenever the vehicle k crosses an urban cell having a roadside unit, then the vehicle k is covered.

V. Baselines

The FPF strategy demands partial mobility information, i.e., the concentration of vehicles within the urban cells and the migration ratios between all pairs of urban cells. FPF does not consider any individual information.

We compare our FPF strategy to two deployment strategies summarizing most of the research efforts conducted over the past years. As we have already mentioned, initial deployment strategies do not assume any mobility information and they allocate the roadside units within the densest locations of the road network. Complementary, modern deployment proposals rely on full knowledge of the vehicles trajectories. Both deployment strategies are investigated by Trullols *et al.* in [6].

A. MCP-g: Deployment Based on the Full Trajectories Info

By assuming we have previous knowledge of the trajectories of all vehicles we may formulate the deployment of roadside units as a Maximum Coverage Problem [6].

Definition 1 (Maximum Coverage Problem): Suppose a collection of sets $S = \{S_1, S_2, \ldots, S_m\}$ defined over a domain of elements $X = \{x_1, x_2, \ldots, x_n\}$. Sets may share elements. The goal is to find a collection of sets $S' \subseteq S$ such that the number of covered elements $| \bigcup_{S_i \in S'} |$ is maximized.

Although the Maximum Coverage Problem is NP-Hard, it is well-known that the greedy heuristic achieves an approximation factor of $(1 - 1/m)^m$, where m is the maximum cardinality of the sets in the optimization domain. Trullols *et al.* named this heuristic MCP-g. In order to cover a given region MCP-g iteratively selects those α urban cells having the largest number of uncovered vehicles.

Algorithm 2 MCP-g (Requires the Trajectories of all Vehicles)

Input: G, α, ϕ;
Output: Γ (locations to receive RSUs);
1: $\Gamma \leftarrow \emptyset$; ▷ solution set starts empty
2: **for** $i = 1$ to α **do**
3: $M \leftarrow$ UpdateFlow(G, ϕ);
4: $\varphi \leftarrow$ GetCellPresentingMaxVehicles(M, G);
5: $\Gamma \leftarrow \Gamma \cup \varphi$;
6: $M \leftarrow M - \varphi$;
7: $G \leftarrow$ RemoveVehiclesCrossingUrbanCell(φ, G)
8: **end for**
9: **return** Γ;

MCP-g receives as input the trajectories information G, the number of available roadside units α, and the grid setup ϕ and selects sets (i.e., urban cells) according to one rule: at each stage, choose the set containing the largest number of uncovered elements (in-depth discussion of MCP-g is presented in [6]).

The Algorithm 2 presents a naive implementation of MCP-g. At each iteration we update the bi-dimensional matrix M with the corresponding number of uncovered vehicles crossing each urban cell using the trajectories information G and the grid setup ϕ. Then, we select the urban cell presenting the largest amount of vehicles φ. Then, we add the urban cell φ to the solution set Γ. After that, we remove φ from set of candidates urban cells M, and we remove all vehicles crossing φ from the mobility trace G incurring in a computational complexity of $\Theta(\alpha.M.G)$.

Throughout this text we assume that MCP-g offers a close-to-optimal solution when considering the Cologne scenario: in order to evaluate the performance of MCP-g we have implemented the ILPF presented from Equations (3) to (5) using the IBM CPLEX Optimizer. The ILPF reaches 93.3% of all vehicles when covering 1% of the Cologne's road network. For a matter of comparison, MCP-g yields 91.9% for the same measure.

B. MCP-kp: Deployment not Based on Mobility Info

Tracking individual vehicles raises several privacy (and practical) issues, which have motivated us to perform this investigation. When we don't have the trajectories information, we cannot solve the Maximum Coverage Problem. All we can do is select the most crowded urban cells in the hope to get an approximate solution of the Maximum Coverage Problem.

Trullols *et al.* call this heuristic 0-1 Knapsack Problem (kp). We use MCP-kp as an approximate solution of the Maximum Coverage Problem when we don't have the vehicles trajectories. The Algorithm 3 presents MCP-kp.

Algorithm 3 MCP-kp (Does not Consider Mobility Info)

Input: M, α;
Output: Γ (solution set);
1: $\Gamma \leftarrow \emptyset$;
2: **for** $i = 1$ to α **do**
3: $\varphi \leftarrow$ GetCellPresentingMaxVehicles(M);
4: $\Gamma \leftarrow \Gamma \cup \varphi$;
5: $M \leftarrow M - \varphi$;
6: **end for**
7: **return** Γ;

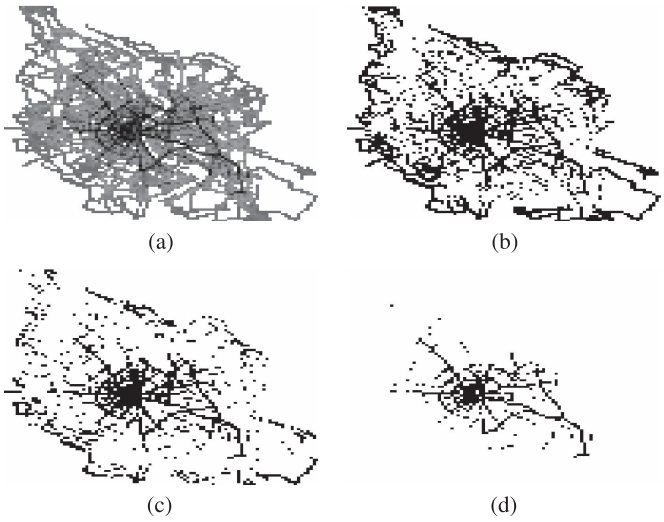


Fig. 3. Figure illustrates merging flows converging to attraction areas in the realistic mobility trace of Cologne, Germany. Roads with high level of traffic are typically interconnected. Such interconnection must be taken into account when we intend to minimize the number of RSUs covering a road network for the dissemination of small traffic announcements. (a) Original flow. (b) Threshold: 40%. (c) Threshold: 60%. (d) Threshold: 80%.

MCP-kp receives the matrix of urban cells M, and the number of available roadside units α to be deployed. Each $M_{i,j}$ stores the number of vehicles crossing the urban cell $\{i, j\}$. MCP-kp iteratively selects the α-densest urban cells. Notice that MCP-kp does not exploit any mobility information at all. In order to select the densest α urban cells MCP-kp incurs in a computational cost of $\Theta(\alpha.M)$.

Although placing the roadside units at the densest places may seem reasonable at a first glance, the assumption fails when we consider that vehicles composing such dense regions are originated from nearby, and the dense region is created as a result of merging flows. Fig. 3 presents the realistic flow of vehicles in Cologne, Germany[5] and we notice flows creating dense areas and some level of connectivity among the densest areas (the darkest the area, the more intense is the flow). The next section presents experiments demonstrating that placing the roadside units at densest locations incurs in a poor deployment performance.

VI. Evaluation Considering a Realistic Scenario

Now we present a set of experiments comparing the performance of FPF, MCP-g, and MCP-kp considering the realistic mobility trace of Cologne, Germany.[6] The trace is composed of (approximately) 10 000 s of traffic and a total of 75 515 vehicles (vehicles join and leave the simulation realistically). Our goal is to quantify the impact of the mobility information (none, partial, and full) on the deployment performance. Recall that MCP-kp lacks any mobility information, but simply sorts the urban cells and selects those α-densest urban cells.

On the other hand, FPF considers migration ratios between urban cells: FPF starts by picking the most crowded urban cell. Then, it projects the flow according to the stochastic matrix

[5]http://kolntrace.project.citi-lab.fr/

[6]http://kolntrace.project.citi-lab.fr/

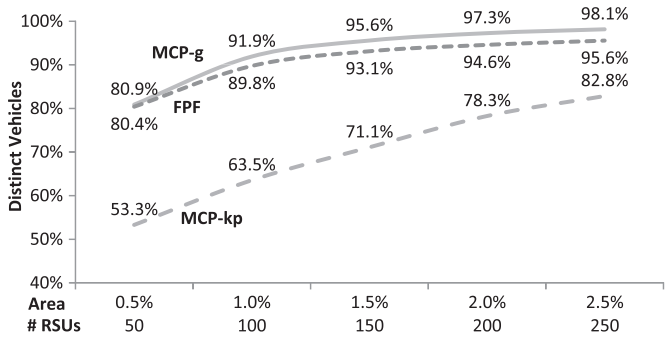


Fig. 4. Percentage of distinct covered vehicles $\times$ covered area. The x-axis shows the number of deployed roadside units. The y-axis indicates the percentage of vehicles experiencing at least one V2I contact opportunity.

of migration ratios (P) and selects the urban cell presenting the highest expectancy of vehicles. Then, it projects the flow and selects another urban cell during $\alpha - 1$ iterations. MCP-g considers full knowledge of the trajectories of the vehicles and always selects the urban cells containing the highest number of uncovered vehicles.

All experiments are performed using the SUMO[7] simulator and a set of tools designed by our team. SUMO runs the mobility trace and outputs the location of each vehicle (our mobility trace T) over time. Then, we compute the bounding box of the mobility trace and partition into a grid of $\phi \times \phi$ urban cells.

We partition the road network into 100×100 grid ($\phi = 100$) in order to achieve urban cells able to be potentially covered by just one roadside unit. A 100×100 grid partitions the Cologne's road network into urban cells of approximately 270 m $\times$ 260 m, a typical range assumed to present high contact probability according to real measurements presented in [27].

We validated our programs by implementing the Integer Linear Programming Formulation presented in Equations (3)–(5) using the IBM CPLEX Optimizer.

A. Covered Vehicles $\times$ Covered Area

Fig. 4 presents the percentage of vehicles experiencing at least one V2I contact opportunity when we deploy the roadside units according to FPF, MCP-kp, or MCP-g. The x-axis shows the percentage of the covered area (and the number of deployed roadside units), while the y-axis indicates the percentage of distinct-and-covered vehicles.

By covering 1.0% of Cologne (selection of 100 urban cells out of 10 000), we reach the following percentage of vehicles: MCP-g $= 91.9\%$; FPF $= 89.8\%$; MCP-kp $= 63.5\%$.

FPF improves MCP-kp in 41.3%. Recall that the only distinction between FPF and MCP-kp are the migration ratios. Complementary, MCP-g uses the trajectories of the vehicles in order to improve FPF in just 2.3%.

Such result demonstrates that previous knowledge of the trajectories of the vehicles are not mandatory for achieving a close-to-optimal deployment performance when we intend to disseminate small traffic announcements. Furthermore, when we increase the covered area in 100%, we improve MCP-g in $+5.8\%$, FPF in $+5.4\%$, and MCP-kp in $+23.1\%$.

[7]Sumo Simulator: http://sumo-sim.org

TABLE I
COVERED AREA × COVERED VEHICLES

	Scen. #1	Scen. #2	Improvement
Covered Area	1.0%	2.0%	100%
Deployed Roadside Units	100	200	100%
MCP-g	91.9%	97.3%	+5.8%
FPF	89.8%	94.6%	+5.4%
MCP-kp	63.5%	78.3%	+23.1%

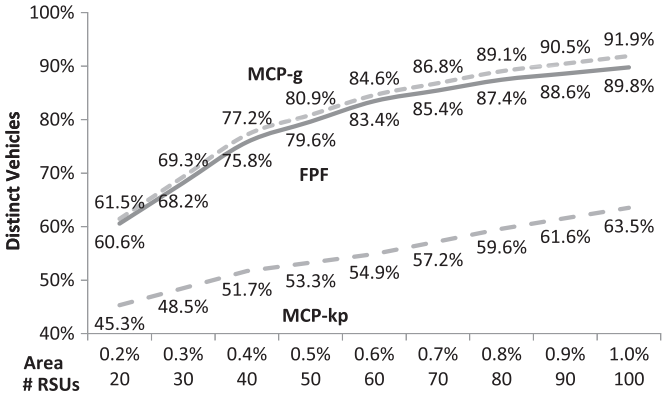


Fig. 5. Low scale deployment. The x-axis indicates the number of deployed roadside units. The y-axis indicates the percentage of vehicles experiencing at least one V2I contact opportunity.

Table I presents the measurements.

Considering that the coverage improvement by adding more than 100 roadside units is too small, we perform the following experiments considering the deployment of just 100 roadside units in order to prevent biases caused by over-deploying roadside units.

B. Low Scale Deployment

In this experiment we analyze the infrastructure performance for a low-scale-incremental deployment.

Fig. 5 presents the deployment performance considering small increments of 10 newly deployed roadside units. FPF and MCP-g show almost the same performance in terms of distinct vehicles reaching the infrastructure, and MCP-kp shows a poor performance. Such issue quantitatively demonstrates that placing the roadside units at the densest locations of the road network (without taking the mobility into account) is far from optimal.

C. Coverage Improvement Analysis

Now we highlight the percentage improvement in terms of the number of distinct-and-covered vehicles for pairs of strategies: Fig. 6 plots the improvement of **FPF over MCP-kp** (Relation i_1), and the improvement of **MCP-g over FPF** (Relation i_2).

- Relation i_1 is given by $(\text{FPF}/\text{MCP}_{\text{kp}}) - 1$;
- Relation i_2 is $(\text{MCP}_{\text{g}}/\text{FPF}) - 1$.

The relation i_1 peaks at 50.9% when the x-axis is 0.5%. The relation i_2 does not peak. When we increase the number of roadside units, MCP-kp gradually approximates to FPF and MCP-g: beyond 100 roadside units both FPF and MCP-g become saturated and the remaining uncovered vehicles are hard to be captured because they are sparsely distributed.

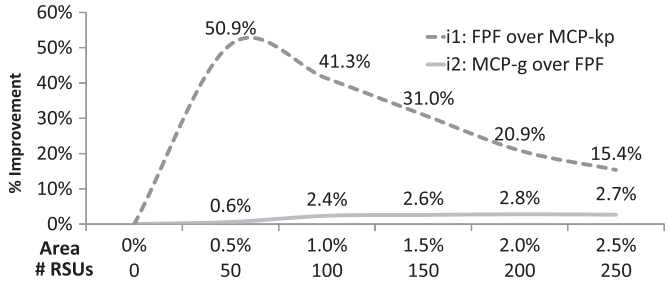


Fig. 6. Percentage improvement of MCP-g over FPF, and FPF over MCP-kp in terms of the trips experiencing at least one V2I contact opportunity. The x-axis indicates the number of deployed roadside units. The y-axis indicates the percentage improvement.

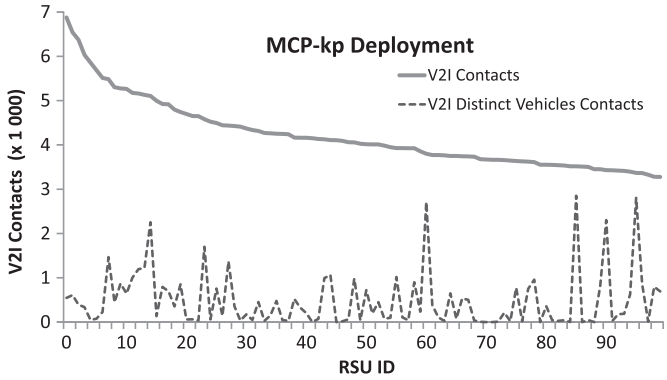


Fig. 7. MCP-kp: Number of V2I contacts per RSU. The x-axis presents the RSUs in the order of deployment. The y-axis indicates the total number of V2I contacts for each RSU. The plot shows the total V2I contacts (continuous line) and the number of V2I distinct vehicles contacts for a given RSU.

D. Number of V2I Contacts per Roadside Unit

Now we characterize the total number of V2I contact opportunities (one vehicle may contact the infrastructure more than once during the trip).

Fig. 7 plots the number of V2I contact opportunities per roadside unit considering the MCP-kp deployment. The x-axis indicates the ID of the roadside unit (ascending from 0 up to 99 according to the order of the deployment), while the y-axis indicates the number of V2I contact opportunities.

The continuous line indicates the total number of V2I contacts for each roadside unit. Since MCP-kp selects the densest urban cells for deploying the roadside units, the number of V2I contacts decreases as we increase the roadside units' ID (the first roadside unit is placed at the densest urban cell, the second roadside unit is placed at the second densest urban cell, and so forth).

The dotted line indicates the number of distinct vehicles contacting the infrastructure (notice that we consider just the first contact opportunity of each vehicle during the entire simulation). Some roadside units are responsible for contacting a high number of vehicles for the first time. On the other hand, other roadside units never contact any vehicle for the first time (e.g., RSU #70). Each roadside unit covers a median of 220 ± 122 vehicles considering a confidence interval of 95%.

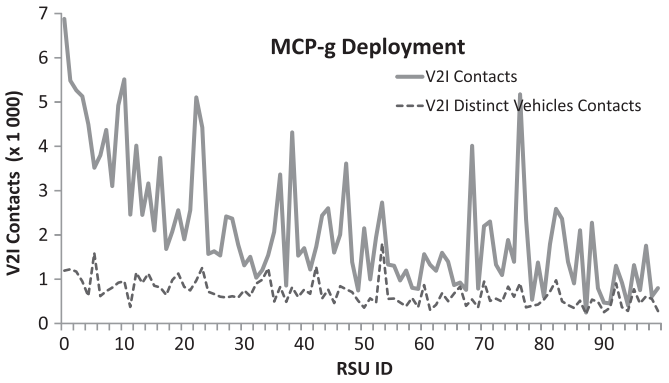


Fig. 8. MCP-g: Number of V2I contacts per RSU. The x-axis presents the RSUs in the order of deployment. The y-axis indicates the total number of V2I contacts for each RSU. Plot shows the total V2I contacts (continuous line) and the number of V2I distinct vehicles contacts for a given RSU (dotted).

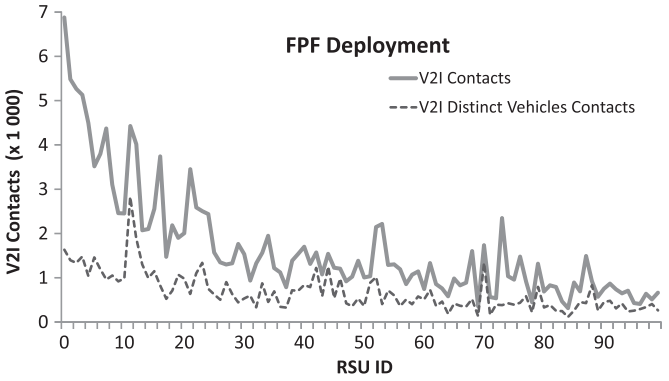


Fig. 9. FPF: Number of V2I contacts per RSU. The x-axis presents the RSUs in the order of deployment. The y-axis indicates the total number of V2I contacts for each RSU. Plot shows the total V2I contacts (continuous line) and the number of V2I distinct vehicles contacts for a given RSU (dotted).

Fig. 8 presents the same analysis, but now considering the MCP-g deployment. The continuous line indicates that the number of total V2I contacts presents a high variance for distinct roadside units. On the other hand, the number of distinct vehicles contacting the roadside units (dotted line) presents less variance than MCP-kp, after all, MCP-g selects the urban cells according to the number of uncovered vehicles. Each roadside unit covers a median of 619 ± 57 vehicles considering a confidence interval of 95%.

Fig. 9 presents the same analysis, but now considering a FPF deployment. The dotted line indicates a greedy behavior for FPF. Each roadside unit covers a median of 532 ± 84 vehicles considering a confidence interval of 95% (Table II).

This experiment demonstrates that relying on individual vehicles trajectories provides a small increase on the deployment performance (from 677 ± 84 to 694 ± 57 distinct vehicles per roadside unit). When we consider the total number of contacts, MCP-kp achieves 796 021 contacts, followed by MCP-g with 371 647 (-53.3%), and FPF with 222 049 (-72.1%). FPF reduces the contact opportunities in 72.1% when compared to MCP-kp. On the other hand, FPF increases the number of distinct-and-covered vehicles in 26.3%. Table II summarizes the results.

TABLE II
DISTINCT-AND-COVERED VEHICLES PER ROADSIDE UNIT

	Average	Median	Confidence (95%)
MCP-kp	486	220	122
FPF	677	532	84
MCP-g	694	619	57

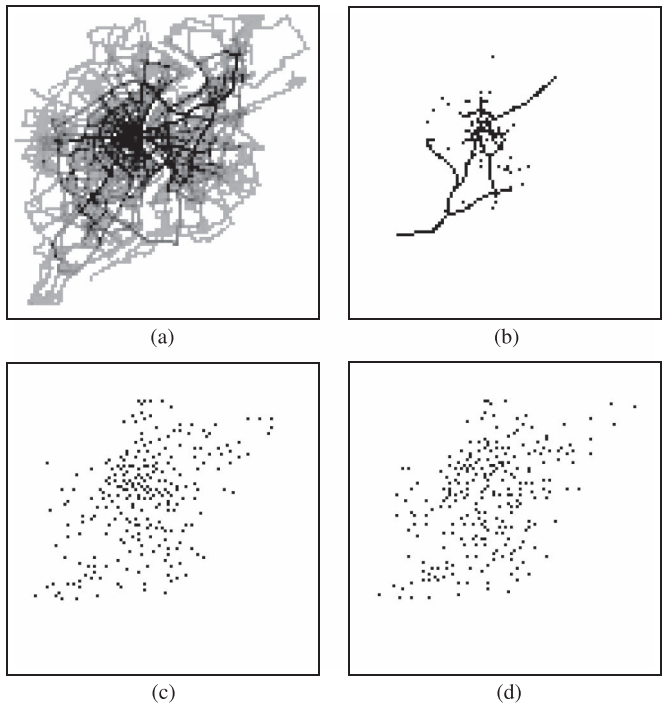


Fig. 10. Road network and layout of roadside units for: MCP-kp, MCP-g, and FPF. (a) Road network. (b) MCP-kp. (c) MCP-g. (d) FPF.

E. Roadside Units Layout

Here we present the layout of the roadside units according to each strategy.

The routes are illustrated in Fig. 10(a) (the darkest the area, the more intense is the flow). The MCP-kp deployment is indicated in Fig. 10(b) (black dots represents roadside units). As we can notice, the roadside units are allocated at very popular urban cells. Those vehicles traveling these very popular routes experience several contact opportunities. On the other hand, vehicles out of these popular routes do not receive any V2I contact opportunity at all.

The MCP-g deployment is presented in Fig. 10(c), a very distinct layout from MCP-kp. MCP-g deploys the roadside units in order to reach the maximum number of distinct vehicles, so the roadside units are distributed along the entire road network. Finally, the FPF deployment is presented in Fig. 10(d), and we notice a high similarity to MCP-g.

F. Number of Roadside Units Crossed per Vehicle

As we have previously mentioned, the goal of our strategy is to reduce the redundant coverage in order to better distribute the contact opportunities. Therefore, in Fig. 11 we present the number of roadside units crossed per vehicle considering the three deployment strategies. The x-axis indicates the number of crossed roadside units. The y-axis indicates the number of

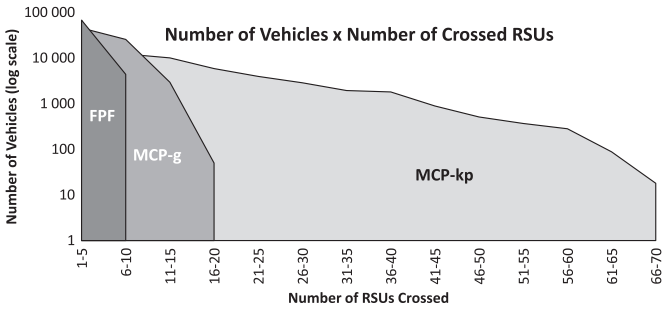


Fig. 11. Number of crossed RSUs per vehicle. The x-axis indicates the amount of RSUs crossed. The y-axis indicates the number of vehicles $(\log_{10})$.

vehicles crossing that amount of roadside units (notice we use a logarithmic scale in the y-axis).

- Vehicles from FPF reach up to 10 roadside units.
- Vehicles from MCP-g reach up to 20 roadside units.
- Vehicles from MCP-kp reach up to 70 roadside units.

MCP-kp presents more vehicles driving through roadside units because it concentrates roadside units at very popular locations, and vehicles crossing such popular locations experience several contact opportunities. The strategy to concentrate roadside units at very popular locations has a side-effect: 36.5% of the MCP-kp vehicles never cross any roadside unit. For a matter of comparison, the same measure for FPF is 10.2%, while for MCP-g it is just 8.1%.

Thus, FPF is able to reduce the redundant coverage provided by MCP-kp. On the other hand, MCP-g provides a better distribution of contact opportunities at the cost of collecting and processing the trajectories of all vehicles.

G. Mobile Roadside Units for Traffic Fluctuations

The Cologne mobility trace has data from 6:00 am up to 8:40 am. Since we are dealing with a realistic trace, the traffic fluctuates along this time period. In order to identify the possible gains when using an architecture combining both stationary roadside units and mobile roadside units (e.g., drones), we investigate the network performance by recomputing the location of each roadside unit every 20 min using FPF. The interval of 20 min is arbitrary. The only goal of this experiment is to investigate whether such an hybrid architecture would improve the network performance.

Thus, now we are assuming the migration ratios and the density of vehicles computed for the last 20 min-window. In Fig. 12 we present the improvement in terms of the number of distinct vehicles experiencing V2I contact opportunities. The x-axis indicates the time interval, while the y-axis indicates the improvements on the number of distinct vehicles (compared to the previous time window).

In addition, the Fig. 13 presents the percentage of roadside units changing locations in order to address the traffic fluctuation. At 6:20 am 58% of the roadside units change their locations when compared to the layout of 6:00 am indicating a high variance of the flow. However, at 8:40 am only 36% of the roadside units move to a new location, indicating a more stabilized traffic flow.

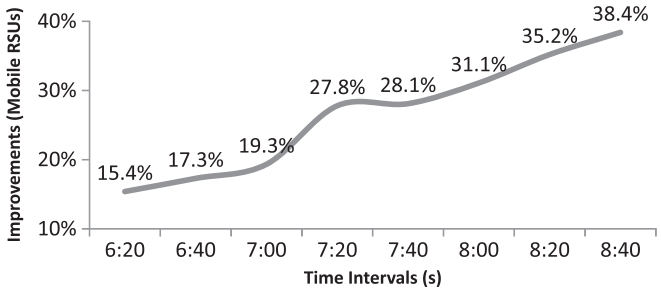


Fig. 12. Increase on V2I contacts when moving the roadside units every 20 min. The x-axis indicates the time interval. The y-axis indicates the improvements on the number of distinct vehicles.

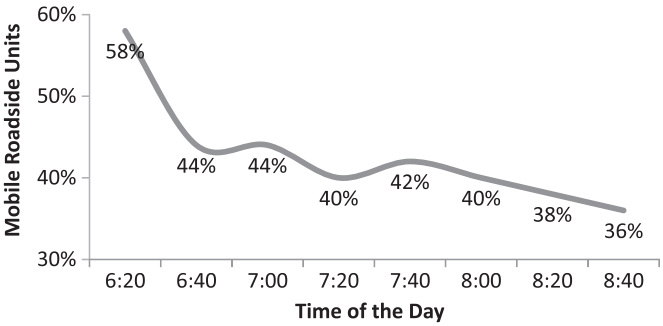


Fig. 13. Percentage of roadside units moving in order to address the traffic changes.

VII. Conclusion

In this work we propose a deployment algorithm based on migration ratios between urban cells without relying on the individual vehicles trajectories. We compare our approach to MCP-kp and MCP-g [6]. MCP-g is the greedy heuristic of the Maximum Coverage Problem and it requires full knowledge of vehicles trajectories (full mobility information). On the other hand, MCP-kp deploys the roadside units at the densest locations of the road network.

We consider the realistic trace of Cologne, Germany, and our results demonstrate that previous knowledge of the vehicles trajectories are not mandatory for achieving a close-to-optimal deployment performance when we intend to disseminate small traffic announcements. Finally, our FPF provides a deployment layout very similar to the one obtained when considering a full mobility information.

References

[1] H. Hartenstein and K. Laberteaux, "A tutorial survey on vehicular *ad hoc* networks," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 164–171, Jun. 2008.
[2] J. Blum, A. Eskandarian, and L. Hoffman, "Challenges of intervehicle *ad hoc* networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 347–351, Dec. 2004.
[3] A. Reis, S. Sargento, and O. Tonguz, "On the performance of sparse vehicular networks with road side units," in *Proc. IEEE 73rd VTC—Spring*, May 2011, pp. 1–5.
[4] K. Mershad, H. Artail, and M. Gerla, "Roamer: Roadside units as message routers in VANETs," *Ad Hoc Netw.*, vol. 10, no. 3, pp. 479–496, May 2012.
[5] R. H. Frenkiel, B. Badrinath, J. Borres, and R. D. Yates, "The infostations challenge: Balancing cost and ubiquity in delivering wireless data," *IEEE Pers. Commun.*, vol. 7, no. 2, pp. 66–71, Apr. 2000.
[6] O. Trullols, M. Fiore, C. Casetti, C. Chiasserini, and J. B. Ordinas, "Planning roadside infrastructure for information dissemination in intelligent transportation systems," *Comput. Commun.*, vol. 33, no. 4, pp. 432–442, Mar. 2010.

[7] C. M. Silva, A. L. Aquino, and W. Meira, Jr., "Deployment of roadside units based on partial mobility information," *Comput. Commun.*, vol. 60, pp. 28–39, Apr. 2015.
[8] M. Nekoui, A. Eslami, and H. Pishro-Nik, "The capacity of vehicular *ad hoc* networks with infrastructure," in *Proc. IEEE 6th Int. Symp. WiOPT*, Apr. 2008, pp. 267–272.
[9] Z. Zheng, P. Sinha, and S. Kumar, "Alpha coverage: Bounding the interconnection gap for vehicular internet access," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 2831–2835.
[10] P. Li, X. Huang, Y. Fang, and P. Lin, "Optimal placement of gateways in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3421–3430, Nov. 2007.
[11] Z. Zheng, Z. Lu, P. Sinha, and S. Kumar, "Maximizing the contact opportunity for vehicular internet access," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.
[12] J. Lee and C. Kim, "A roadside unit placement scheme for vehicular telematics networks," in *Advances in Computer Science and Information Technology*, vol. 6059, ser. Lecture Notes in Computer Science T.-H. Kim and H. Adeli, Eds. Berlin, Germany: Springer-Verlag, 2010, pp. 196–202.
[13] J. Chi, Y. Jo, H. Park, and S. Park, "Intersection-priority based optimal RSU allocation for VANET," in *Proc. IEEE 5th ICUFN*, Jul. 2013, pp. 350–355.
[14] Y. Liang, H. Liu, and D. Rajan, "Optimal placement and configuration of roadside units in vehicular networks," in *Proc. IEEE 75th VTC—Spring*, May 2012, pp. 1–6.
[15] Y. Xiong, J. Ma, W. Wang, and D. Tu, "Roadgate: Mobility-centric roadside units deployment for vehicular networks," *Int. J. Distrib. Sens. Netw.*, vol. 2013, 2013, Art. ID 690974.
[16] C. M. Silva and W. Meira, "Planning the roadside communication infrastructure for vehicular networks with QoS guarantees," in *Proc. IEEE ISCC*, Jul. 2015.
[17] C. M. Silva and W. Meira, "Evaluating the performance of heterogeneous vehicular networks," in *Proc. IEEE VTC*, Sep. 2015.
[18] B. Aslam, F. Amjad, and C. Zou, "Optimal roadside units placement in urban areas for vehicular networks," in *Proc. IEEE ISCC*, Jul. 2012, pp. 423–429.
[19] S. Habib and M. Safar, "Sensitivity study of sensors' coverage within wireless sensor networks," in *Proc. IEEE 16th ICCCN*, 2007, pp. 876–881.
[20] N. Lu *et al.*, "Vehicles meet infrastructure: Toward capacity-cost tradeoffs for vehicular access networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1266–1277, Sep. 2013.
[21] B. Xie, G. Xia, Y. Chen, and M. Xu, "Roadside infrastructure placement for information dissemination in urban its based on a probabilistic model," in *Network and Parallel Computing*, vol. 8147, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2013, pp. 322–331.
[22] C. M. Silva, A. L. Aquino, and W. Meira, "Design of roadside infrastructure for information dissemination in vehicular networks," in *Proc. IEEE NOMS*, May 2014, pp. 1–8.
[23] O. Tonguz and W. Viriyasitavat, "Cars as roadside units: A self-organizing network solution," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 112–120, Dec. 2013.
[24] M. Jerbi, S. Senouci, Y. Doudane, and A. L. Beylot, "Geo-localized virtual infrastructure for urban vehicular networks," in *Proc. IEEE 8th Int. Conf. ITST*, Oct. 2008, pp. 305–310.
[25] J. Luo, X. Gu, T. Zhao, and W. Yan, "MI-VANET: A new mobile infrastructure based VANET architecture for urban environment," in *Proc. IEEE 72nd VTC—Fall*, Sep. 2010, pp. 1–5.
[26] C. M. Silva, "Deployment of roadside units based on partial mobility information," Dept. de Ciência da Computação da Univ. Fed. de Minas Gerais, Belo Horizonte, Brazil, Dec. 2014. [Online]. Available: http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/ESBF-9TELS9
[27] F. Teixeira, V. Silva, J. Leoni, D. Macedo, and J. M. S. Nogueira, "Vehicular networks using the ieee 802.11p standard: An experimental analysis," *Veh. Commun.*, vol. 1, no. 2, pp. 91–96, Apr. 2014.

Cristiano M. Silva received the B.S. and M.S. degrees in computer science from the Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, in 2000 and 2004, respectively, the M.B.A. degree from the IBMEC, Rio de Janeiro, Brazil, in 2008, the Specialization in Finances from the Universidade Cândido Mendes, Rio de Janeiro, in 2010, and the Ph.D. degree in computer science from the Universidade Federal de Minas Gerais, in 2014. He is currently an Associate Professor at the Universidade Federal de São João del Rei, São João del Rei, Brazil. Dr. Silva serves as a General Chair for the First IEEE/IFIP International Workshop on Urban Mobility and Intelligent Transportation Systems (UMITS 2016) co-located with IEEE NOMS. He was a recipient of an award at the National Contest of Scientific Initiation, the Best Ph.D. Thesis in Computer Science at the Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, and the Best Paper Award in two consecutive years at the Brazilian Symposium on Ubiquitous and Pervasive Computing (SBCUP'2013–2014), an event promoted by the Brazilian Computing Society.

Wagner Meira, Jr., received the Ph.D. degree from the University of Rochester, Rochester, NY, USA, in 1997. He is a Full Professor at the Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. He has published more than 200 papers in top venues and is a coauthor of the book *Data Mining and Analysis—Fundamental Concepts and Algorithms* (Cambridge Univ. Press, 2014). His research focuses on scalability and efficiency of large-scale parallel and distributed systems, from massively parallel to Internet-based platforms, and on data mining algorithms, their parallelization, and application to areas such as information retrieval, bioinformatics, and e-governance.

João F. M. Sarubbi received the B.S. degree in computer science from the Universidade Estadual do Ceará, Fortaleza, Brazil, in 2001 and the M.S. and Ph.D. degrees in computer science from the Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, in 2003 and 2008, respectively. He is currently an Associate Professor at the Departamento de Computação, Centro Federal de Educação Tecnológica de Minas Gerais, Divinópolis, Brazil. His research focuses on optimization, mathematical programming, and intelligent transportation systems.