Improving information dissemination in vehicular networks by selecting appropriate disseminators

Fabrício A. Silva*†, Thais R. M. B. Silva†, Fabrício Benevenuto*, Linnyer B. Ruiz‡ and Antonio A. F. Loureiro*
*Computer Science Department - Federal University of Minas Gerais
†Florestal Campus - Federal University of Viçosa
‡Department of Informatics - State University of Maringá
Emails: {fabricio.asilva, thaisrb, fabricio, loureiro}@dcc.ufmg.br, {linnyer}@gmail.com

***Abstract*—Information dissemination applications (e.g. weather, news, advertisements, traffic, and so on) are among the most promising ones for vehicular networks. However, disseminating information to all vehicles in vehicular networks is not a trivial task because of their specific characteristics like high dynamic topology and different density values along time and space. The proposals found in literature focused on improved flooding schemes that, in realistic scenarios, may cause unnecessary overhead and redundant data. This work addresses the dissemination to all vehicles problem by assuming that some vehicles are more appropriate to be good disseminators than others. This hypothesis was validated through a large-scale urban mobility scenario and two heuristics were proposed to select the appropriate disseminators. Results showed that the proposed heuristics were able to achieve high coverage (more than 97% of vehicles) with low redundant data.**

I. INTRODUCTION

Advances in wireless communications and embedded systems observed in the last years allowed a new network technology to emerge: vehicular networks. In this kind of network, vehicles are able to communicate wirelessly with each other as well as with infrastructure stations.

According to Willke et al. [1], information system applications that disseminate information (e.g. traffic, weather, advertisements, news, and so on) to a number of vehicles are among the most promising applications for vehicular networks. Several challenges arise when the information must be disseminated to all vehicles in the network, like coverage, redundancy, scalability and communication issues.

Specific characteristics of vehicular networks turn the dissemination process one of its greatest challenges [2]. First, the high mobility of vehicles causes constant changes in the network topology. In addition, transmissor and receiver may not be active simultaneously or in direct contact with each other. Finally, the network density may vary spatial and temporally, which causes disconnections on sparse areas and intense data traffic on dense ones. In applications in which the whole network coverage is a requirement (e.g. information must be disseminated to all vehicles), these characteristics become even more critical.

Flooding is the most adopted strategy to disseminate information to all vehicles [3]. In this strategy, vehicles receive, store and forward messages to its neighbors, which guarantee with high probability that all vehicles will be reached. However, unnecessary overhead may arise even when improved flooding techniques are adopted. Furthermore, collisions and congestions may occur as redundant messages will be exchanged.

Given all these issues, one question arises: what if, despite all vehicles, we select just some of them to act as disseminators? If this is a good option, which vehicles' characteristics should be considered in such selection process? In this work, we validate the hypothesis that is probable a good idea to select some vehicles (vehicles that are more relevant to the network) to perform the dissemination. Furthermore, we present the network characteristics that are important to help on such selection. In addition to this, we also propose and evaluate two heuristics to select good disseminators. The proposed heuristics were compared to other strategies, including an improved version of flooding, and the results demonstrate their efficacy regarding network coverage and redundancy.

The remainder of this paper is organized as follows. Section II presents the related work. Section III comprises a realistic urban vehicular mobility trace characterization and the validation of the hypothesis that some vehicles are more relevant and would be good disseminators, which is the first contribution of this work. In the following, the heuristics to select good vehicles to act as disseminators are described in Section IV and the evaluation process and results are presented in Section V. This is the second contribution of this work. Finally, our concluding remarks and future work are listed in Section VI.

II. RELATED WORK

Several papers found in literature address the dissemination problem in vehicular networks. Given the increasing interest in this field, some researchers have already published surveys like [1], [2], [3]. Most of the authors propose improvements to the pure flooding scheme with the goal of diminishing its overhead with no loss in coverage. These solutions generally adopt selective forwarding rules based on vehicles' position, direction, and other information. Some of these works are briefly described below.

Rostamzadeh et al. [4] define a selective forwarding scheme in which vehicles wait for their neighbors before forwarding a

This work was financially supported by the Brazilian National Council for Scientific and Technological Development (CNPq) research agency, grant no. 573.738/2008-4 (INCT NAMITEC)

message to avoid duplications in the vicinity, and each vehicle forwards a message only once. Ruiz et al. [5] present a dissemination solution based on a tree structure built dynamically. Differently, Sung et al. [6] try to improve the flooding scheme in crossroads areas where forwarding decisions become more important as there are different directions that vehicles could take. Ciccarese et al. [7] present an algorithm that adopts a distribution contention scheme based on RTS/CTS messages exchanged. Villas et al. [8] propose an improved flooding algorithm that selects vehicles to perform dissemination based on their position on preference zones. Furthermore, vehicles adopt a contention time scheme before retransmiting a message. Finally, Viriyasitavat et al. [9] propose a protocol in which vehicles closer to the border of a region of interest are considered more able to achieve yet unachieved vehicles and are then responsible to perform forwarding.

Although interesting ideas, these works were evaluated considering highway or simplified urban scenarios and did not take into account the vehicular mobility characteristics, which may not guarantee that they would perform well on real mobility scenarios.

The proposal of efficient dissemination protocols requires knowledge regarding the network topology characteristics, specially in high dynamic scenarios like vehicular networks. Some authors have already studied the behavior of mobility traces. Xia et al. [10] characterize a mobility trace composed of 12,096 taxis from Beijing during a week, focusing on contact and clustering characteristics. Ahmed et al. [11] analyze 1,200 buses from Seattle in regarding to contact duration, inter-contact intervals and clustering characteristics. In [12], simplified urban (4Km^2) and highway (25Km) scenarios were also characterized. To the best of our knowledge, the only characterization of realistic urban scenario was performed by Uppoor et al. [13]. The authors analysed a trace from Cologne, in Germany, in regarding to vehicles flow and network density. However, to help on developing new dissemination techniques, other characterization aspects of this mobility scenario were studied by us in this work.

Finally, it is also important to mention some studies that address the selection of vehicles to act as disseminators. Resta et al. [14] demonstrate mathematically the delay and capacity restrictions of broadcast dissemination in mobile networks. They showed that it is possible to achieve full capacity only if the best disseminators nodes are correctly selected. This information reinforce our idea that it is possible to select just some vehicles to act as disseminators and yet have good performance. Khabbaz et al. [15] proposed an algorithm in which fixed road-side units (RSU) select a vehicle to deliver a message to other RSU in a highway scenario. In [16], the authors organize nodes in two different groups: the more and less participative in specific regions. They showed that it was possible to achieve a high network coverage by using the more participative nodes. Finally, Ahmed et al. [17] propose to select what they called *hubs* to be responsible of forwarding a message based on the nodes degree (e.g. 10% of nodes with highest degree) on a bus scenario.

Although interesting to the state of the art, most of the works found in literature do not consider realistic and scalable urban mobility scenarios composed of regular vehicles in despite of only buses or taxis. Furthermore, we could not found any work that evaluates network metrics to help on select good disseminators. In this work, we first characterize a realistic vehicular mobility trace and, based on it, we propose and evaluate two heuristics to select good disseminators.

III. CHARACTERIZATION

In this section we present the first contribution of our work which is the characterization of a realistic large-scale urban mobility trace. The results obtained validate the hypothesis that there exist some vehicles with higher potential to be better disseminators. Furthermore, some characteristics that would help on select these vehicles were identified.

A. Data

Among several mobility traces found, the one describe in [18] was chosen to be evaluated in this work since it represents a realistic large-scale urban mobility scenario from Cologne, in Germany. The trace encompasses the mobility of more than 120,000 vehicles from 6:00am to 8:00 with 1 second granularity.

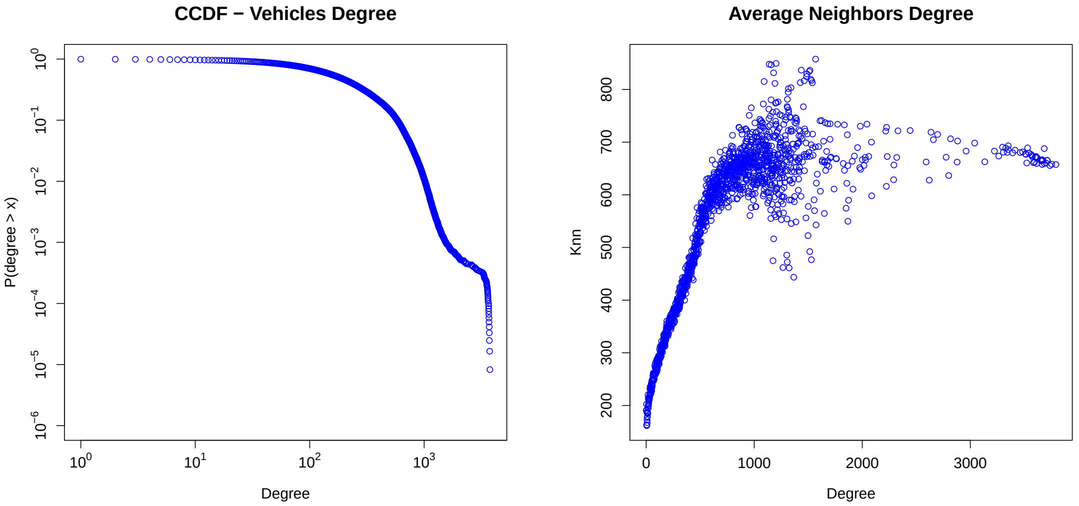
B. Characterization

The network was modeled as a graph $G = (V, E)$ in which V is the vertices set (e.g. vehicles), and $E \subseteq V \times V$ is the set of edges (e.g. contacts among vehicles). An edge $e = (v_i, v_j)$ between vehicles v_i and v_j exists if the distance $d(v_i, v_j)$ between v_i and v_j was lower than 100 meters during a period of time. The value of 100 meters was chosen based on experimental results from [19]. Although other authors have shown that it was possible to achieve a higher distance in vehicular networks, 100 meters was considered a good range for real scenarios with obstacles.

In the following, the characterization results, which include complex networks metrics, are presented.

a) Degree distribution: the degree distribution of vertices is an interesting metric that provides relevant information of the network topology. In this analysis, we considered the complete graph comprised of all vehicles and their contacts during the whole period. This graph comprises 120,913 vehicles and 15,385,919 contacts among them.

Figure 1(a) presents the vehicles degree complementary cumulative distribution function (CCDF) in log-log scale. Most vehicles have a small number of contacts while a small fraction of them have significantly more contacts (minimum=1, median=356 and maximum=3,792). The Maximum Likelihood Estimator (MLE) method was used to check whether this distribution follows a power-law or not, like in other social networks [20]. Considering all values ($x\text{-}min = 1$), the high value of the Kolmogorov-Smirnov statistic (KS = 0.64) indicates that this distribution did not follow a power-law. On the other hand, considering only the tail ($x\text{-}min = 798$), we



(a) Vehicles degree distribution

(b) Knn (Pearson Coefficient: 0,69)

Fig. 1. Vehicles degree distribution follows a power-law on its tail. There is a positive correlation between degree and k_{nn} values

can say that it follows a power-law distribution for $\alpha = 5.88$ (KS = 0.034 in this case).

Besides the degree distribution, we have also evaluated the K_{nn} metric that maps vehicles' degree to their neighbors average degree. As showed in Figure 1(b), the positive correlation between vehicles' degree and K_{nn} values (Pearson coefficient = 0.69) indicates a tendency of vehicles with higher degree also to be in contact with other higher degree ones.

b) Centrality and clustering: centrality and clustering metrics could also give insights regarding the network topology. Figure 2(a) depicts the relationship between vehicles' degree and their correspondent betweenness values. The betweenness of a vertex v_i is the number of minimum paths between any two other vertices that include v_i. The higher the betweenness, the more important to the network the vertex is. We can note a correspondence between the vertex degree and its betweenness. However, there are some vertices with high betweenness and not so high degree.

The closeness is also a centrality metric calculated by the inverse of the sum of all distances between vertex v_i and all other vertices. The higher this value, the lower is the sum of the distances and therefore the vertex is relatively close to the other ones. Figure 2(b) shows that there is a linear correspondence between vertices' degree and closeness since its Pearson coefficient is 0.92.

Another important metric is the clustering coefficient that measures the edges density of a vertex's neighbors. As depicted in Figure 2(c), vertices with lower degree have a higher clustering coefficient in general.

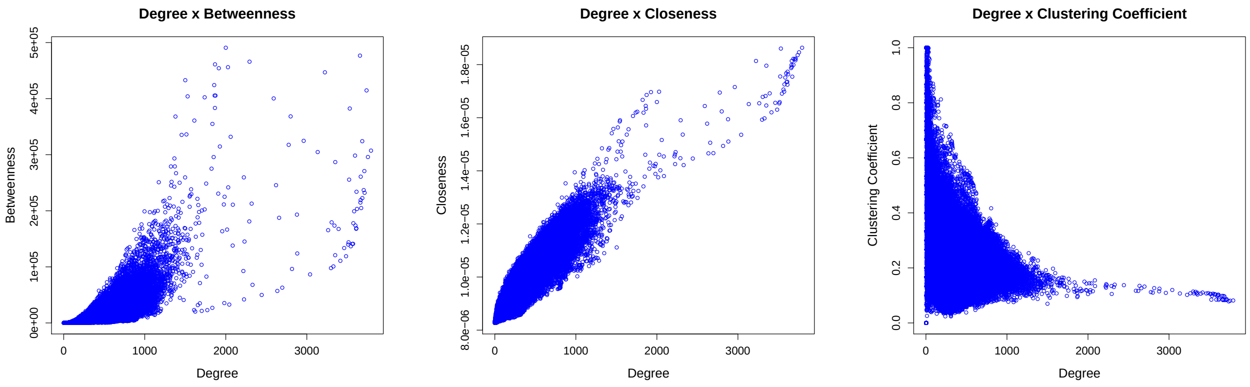
c) Travel time and distance: it is also important to characterize vehicles movement and the contacts among them. Figure 3(a) presents the CCDF of the distance traveled by vehicles and it is possible to note that there are a small number of vehicles that travel for significant higher distances than the majority of the other ones. The same observation is true in regarding to the travel time, as depicted in Figure 3(b), which shows that some vehicles travel for significant longer time than the majority of the other ones. Furthermore, the intersection of the 10% vehicles presenting higher distance traveled with the 10% vehicles with the longer travel time has only 21% of all vehicles. In other words, there are vehicles that travel a long distance for a short period of time, and others that travel short distances for a long time.

We have also analysed how long each contact among vehicles lasted. Figure 3(c) presentes the CCDF of the length of time that each contact lasted and it allows us the realize that most contacts lasted for few seconds, while some of them lasted for much longer time.

C. Hypothesis Validation

Based on the characterization results, it is possible to bring up some insights that validate the idea that some vehicles are more likely to be good disseminators than others, like:

- vehicles' degree distribution indicates that there are some vehicles with significant higher number of contacts than the majority of the other ones;
- centrality analyzes show that some vehicles have higher values of *betweenness* and *closeness*, which indicates that they are more important to the network structure than the others;
- clustering coefficient analysis indicates that some vehicles are more connected to their neighbors of one and two hops than others;
- there are vehicles that travel for longer distances and/or for longer period of times than the majority of others;
- contacts duration analysis indicates that there are connections between vehicles that last for a long period of

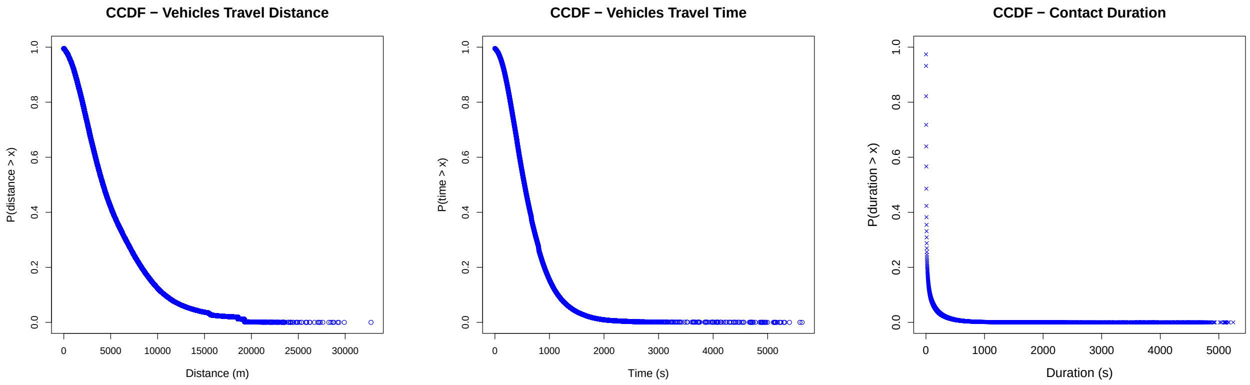


(a) Degree x Betweenness (Pearson Coefficient: 0.74)

(b) Degree x Closeness (Pearson Coefficient: 0.92)

(c) Degree x Clustering Coefficient (Pearson Coefficient: -0.30)

Fig. 2. Centrality and clustering relationships with vertices degree



(a) Distance traveled by vehicles

(b) Travel time for vehicles

(c) Contacts duration

Fig. 3. Movement and contact information. Some vehicles travel long distances and for a long time. Furthermore, some contacts among them lasted for many minutes.

time.

All these remarks reinforce that some vehicles are more important to the network dissemination than others and therefore they may be more appropriate to act as disseminators. However, what features must be considered to select the best disseminators? To answer this question, we evaluated what happens to the network structure when some vehicles with specific characteristics are removed from it.

To this end, let V be the set of all vehicles; D_f the set of $f\%$ vehicles with higher degree; B_f the set of $f\%$ vehicles with higher *betweenness*; C_f the set of $f\%$ vehicles with higher *closeness*; CC_f the set of $f\%$ vehicles with higher clustering coefficient; TT_f the set of $f\%$ vehicles with higher travel time; TD_f the set of $f\%$ vehicles with higher travel distance.

The complete graph G comprises 30 connected components, which are subgraphs in which two vertices are connected to each other by one or more paths and any vertex is not connected to any other in the graph. Table I shows that when 10% of vehicles with specific characteristics are removed from G, the number of connected components increases. It is also possible to note that the clustering coefficient and traveled distance features had a more significant impact as the number of connected components increased from 30 to 81 and 365, respectively.

TABLE I
HOW VEHICLES WITH SPECIFIC CHARACTERISTICS IMPACT ON NETWORK STRUCTURE.

Vehicles Set	**Connected Components**
V	30
$V - D_{10}$	62
$V - B_{10}$	54
$V - C_{10}$	52
$V - CC_{10}$	81
$V - TT_{10}$	60
$V - TD_{10}$	365

Based on the analyses presented, it is possible to conclude that some vehicles have special characteristics that make them more likely to be good disseminators than others and that characteristics like degree, *betweenness*, *closeness*, clustering coefficient, distance traveled and travel time are probably important on the selection of such vehicles. These results are the first contribution of this work. In the next Section, we propose two heuristics to select appropriate disseminators.

IV. HEURISTICS TO SELECT APPROPRIATE DISSEMINATORS

In this section, we describe the second contribution of this work which is the proposal of heuristics to select good disseminators. As showed in the last section, metrics like degree, *betweenness*, *closeness* and clustering coefficient seem to be good alternatives to help on selecting disseminators. However, we need to know ahead the whole network topology. Instead, we propose heuristics that take into account only the distance traveled and traveled time information, which can be estimated by navigation systems embedded in vehicles.

The idea behind the proposed heuristics is to select vehicles that depart from and arrive at different areas in order to cover the highest number of vehicles possible. To this end, the city's map was divided in 500 meters side quadrants. For each quadrant, a fraction f of vehicles that depart from or arrive at it is selected to act as disseminators. From the fraction f, half of it ($f/2$) comes from vehicles that depart from and the other half from vehicles that arrive at each quadrant. In this way, it is more likely that the complete quadrant will be covered since there will be vehicles departing to and arriving at it at different times.

To illustrate, consider a quadrant from which 20 vehicles depart and at which another 20 vehicles arrive. If we are selecting a fraction $f = 0.1 (10\%)$ of vehicles, then $\lceil 0.05 \times 20 \rceil = 1$ will be selected from the departing ones while another one will be selected from the arriving set. Since vehicles depart and arrive at some quadrant, the fraction $f/2$ is used to avoid more vehicles to be selected than expected.

The difference among the heuristics is how vehicles are selected. As presented in Section III-B, there are vehicles that travel longer distances and others that travel for longer times. Then, each heuristic takes into account one of this characteristics:

d) Heuristic 1 (H1): longer travel time: Among the vehicles that depart and arrive at a quadrant, the ones with longer expected travel time are selected.

e) Heuristic 2 (H2): longer travel distance: Among the vehicles that depart and arrive at a quadrant, the ones with longer expected travel distance are selected.

Algorithm 1 presents the procedure of selecting vehicles for both heuristics H1 and H2.

It is important to note that we are assuming that all vehicles have an embedded navigation system that is able to estimate the travel distance and time by knowing the origin and destination points of the vehicle. Furthermore, based on mobility traces studies, like in [21], it is possible to estimate the density of each quadrant. So, the heuristics' inputs are known in advance.

V. EVALUATION AND RESULTS

We evaluated the performance of the proposed heuristics through an analytical model using graph algorithms. It is assumed that only a selected subset $V' \subseteq V$ of all vehicles contains the information to be disseminated to all other vehicles. These selected vehicles $v' \in V'$ are called disseminators and they may receive the information by infra-structured communications. The other vehicles $v \notin V'$ do not have such capabilities or can not afford for such communication cost. Information is disseminated from a vehicle v' to a vehicle v if a direct contact between them occur in any time.

TABLE II

STRATEGIES TO SELECT APPROPRIATE DISSEMINATORS. $f\%$ IS THE FRACTION OF SELECTED VEHICLES, THAT VARIES FROM 1% TO 15% IN THIS WORK.

Strategy	Description
D	V' = {f% of all vehicles with higher degree}
B	V' = {f% of all vehicles with higher *betweenness*}
C	V' = {f% of all vehicles with higher *closeness*}
CC	V' = {f% of all vehicles with higher clustering coefficient}
TD	V' = {f% of all vehicles with longer travel distance}
TT	V' = {f% of all vehicles with longer travel time}

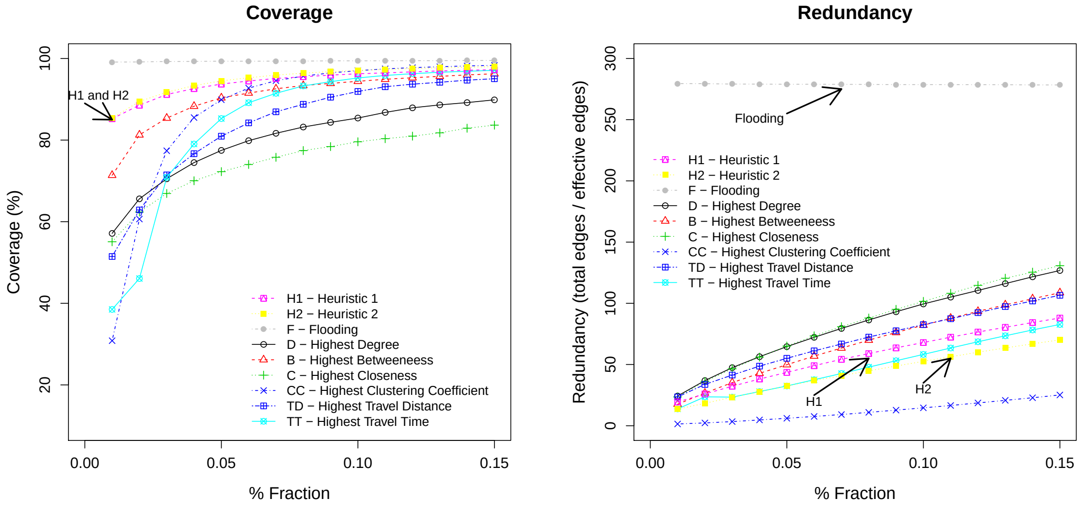
Three metrics were adopted to evaluate the dissemination performance of the heuristics: coverage, redundancy and capacity. Coverage measures the percentual of vehicles that were reached and therefore received the information, being 100% the best result possible. Redundancy stands for the proportional amount of contacts that were unnecessary or duplicated and is computed as the ratio between the total amount of contacts established among the disseminators and the other vehicles and the amount of these contacts that were effectively used. The higher the redundancy value, more duplicated data were disseminated. So, the lower this value, the better. Finally, capacity measures how long all contacts established last in total. The higher this value, the more data could be transmitted and consequently the higher the chance of a communication success.

The proposed heuristics were compared to other 6 strategies as described in Table II. These strategies select the disseminators based on characteristics that were considered relevant after the characterization described in Section III-B. Among all vehicles, fractions varying from 1% to 15% of disseminators were selected, which may help on identifying the upper bound of vehicles to act as disseminators.

We have also compared the heuristics to an improved flooding strategy in which, initially, a fraction of vehicles is selected to be the source of information. In order to avoid collisions and duplications, each vehicle disseminates the information to its neighbors with probability 0.1. As other vehicles receive the information, they store-carry-and-forward it with probability 0.1. This is an improved version of flooding that tries to decrease redundancy. Some studies compute the forwarding probability based on vehicles' information like their position, direction or the area density, to name a few. However, the forwarding probability was estimated as a low value in our evaluation process to represent a good flooding scheme. It is important to state that this strategy was not compromised, since we will show that it achieved approximately 100% of coverage, as expected.

Algorithm 1 Heuristics H1 and H2

```
$f$                                                  ▷ the fraction of vehicles to be selected
$D(v)$                                               ▷ function that returns the travel distance of vehicle $v$
$T(v)$                                               ▷ function that returns the travel time of vehicle $v$
$H1 \leftarrow \{\}$                                 ▷ vehicles selected by H1
$H2 \leftarrow \{\}$                                 ▷ vehicles selected by H2
for all Quadrant q do
    $Q_d \leftarrow \{v_i | v_i\ departs\ from\ q\}$
    $Q_a \leftarrow \{v_i | v_i\ arrives\ at\ q\}$
    $D_d \leftarrow [v_1, v_2, ..., v_{|Q_d|}]$ where $v_i \in Q_d \wedge D(v_i) \geq D(v_j)\ \forall\ i \leq j$
    $D_a \leftarrow [v_1, v_2, ..., v_{|Q_a|}]$ where $v_i \in Q_a \wedge D(v_i) \geq D(v_j)\ \forall\ i \leq j$
    $T_d \leftarrow [v_1, v_2, ..., v_{|Q_d|}]$ where $v_i \in Q_d \wedge T(v_i) \geq T(v_j)\ \forall\ i \leq j$
    $T_a \leftarrow [v_1, v_2, ..., v_{|Q_a|}]$ where $v_i \in Q_a \wedge T(v_i) \geq T(v_j)\ \forall\ i \leq j$
    $M \leftarrow \lceil f/2 \times |Q_d| \rceil$     ▷ number of vehicles that depart from $q$ to be selected
    $K \leftarrow \lceil f/2 \times |Q_a| \rceil$     ▷ number of vehicles that arrive at $q$ to be selected
    $H1 \leftarrow H1 \cup T_d[1...M] \cup T_a[1...K]$    ▷ vehicles with higher travel time
    $H2 \leftarrow H2 \cup D_d[1...M] \cup D_a[1...K]$    ▷ vehicles with higher traveled distance
end for
```



(a) Coverage

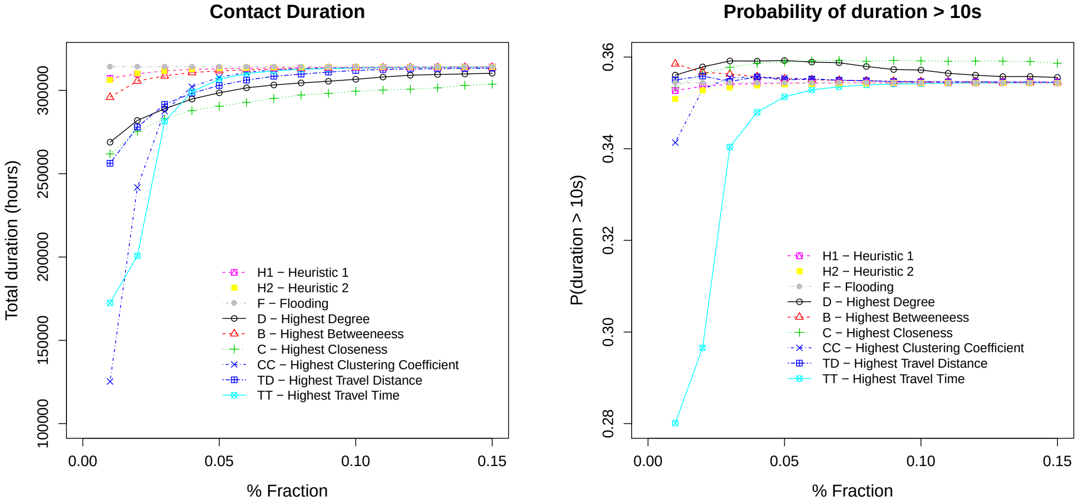
(b) Redundancy

Fig. 4. Both heuristics obtained good coverage and redundancy results. Flooding performs well with regards to coverage, but disseminate much more redundant data.

A. Results

Figure 4 shows the coverage and redundancy results. Regarding coverage (Figure 4(a)), flooding achieved nearly 100% no matter the fraction of disseminators that started the dissemination process. Since dissemination is propagated among vehicles, it is expected that all of them will be reached even with a small forwarding probability. Both heuristics H1 and H2 obtained better results compared to the other strategies for fractions smaller than 6%. From fractions of 7% on, strategies CC, B and TT obtained similar results. As expected, strategies D and C had similar results because these features (degree and closeness) are well correlated as showed in Figure 2(b). It was expected that strategy D could achieve a better result as the more neighbors a vehicle has, the more vehicles it will reach. However, as it is showed by the redundancy results, the vehicles with higher degree have a lot of neighbors in common. Strategies B and CC differ when the fraction is smaller than 5% in which CC performs better. However, from 6% on they had similar results.

Coverage results showed that the proposed heuristics were effective as it was possible to achieve approximately 94% of coverage with only 5% of vehicles selected as disseminators, which is more than any other strategy not including the flooding one. In addition, approximately 97% of coverage was achieved when 11% of vehicles are selected as disseminators. Finally, the results also indicate that the most relevant complex networks characteristics to select good disseminators are clustering coefficient and *betweeness*. However, these



(a) Total Contact Duration

(b) Probability of Long Duration Contacts

Fig. 5. Both heuristics perform well with regards to capacity.

characteristics require the complete graph to be known a priori.

Regarding redundancy, both heuristics obtained good results. Since vehicles are selected based on their departure and arrival positions, the dissemination is expected to be more distributed and consequently generate less redundant data. Heuristic H2 (the one that selects vehicles by their traveled distance) presented lower redundancy values, which indicates that the distance traveled by vehicles is a better metric than their travel time. This occurs probably because vehicles that take longer times may be stuck in traffic jam situations. The flooding strategy presented significantly higher redundancy even with a small forwarding probability of 0.1. This occurs because vehicles have a lot of neighbors in common, which increases the probability of duplications. The clustering coefficient (CC) strategy had good redundancy results because vehicles from different connected components of the network will be selected, diminishing the chance of duplications. Strategies D and C presented the higher redundancy after the flooding one, which indicates that the degree and closeness are not good characteristics to be considered, differently from expected.

The dissemination capacity results, measured by the contact duration, are depicted in Figure 5(a). The same comments regarding the coverage results are valid for this metric, as both heuristics obtained better results until the fraction of 5% and results similar to the other strategies from fraction 6% on. However, flooding was not as good as in coverage for this metric. The four strategies D, C, B and CC had similar results. In other words, any of them would achieve practically the same capacity. However, strategies B and CC obtained better coverage results as showed in Figure 4(a). This indicates that strategies D and C tend to have longer duration contacts, as can be seem in Figure 5(b) which presents the probability of a contact last longer than 10 seconds. In this graphic, it can be seen that strategies D and C have in general higher probability of being part of a long duration contact.

As final remarks, we can state that the objective of the proposed heuristics were achieved as they obtained good coverage and capacity results with lower redundant values. The improved flooding strategy, which is the most adopted in literature, was able to deliver the information to practically 100% of vehicles. However, it originates a significant higher amount of redundant information. Among the other strategies, the one that selects vehicles based on their clustering coefficient achieved coverage results close to the heuristics with a lower number of redundant messages. However, differently from both heuristics, this strategy requires the complete graph to be known a priori.

VI. CONCLUSION AND FUTURE WORKS

In this work we addressed the problem of disseminating information to all vehicles in a vehicular network by following an approach that considers only a selected subset of vehicles to act as disseminators, differently from flooding schemes in which all vehicles perform this action. First, a realistic urban vehicular mobility trace were characterized and the hypothesis that there are vehicles that are more likely to be better disseminators than others was validated. Moreover, some metrics that are representative to help on selecting these vehicles were presented. Then, based on this characterization, we propose two heuristics that select appropriate disseminators and achieved good coverage and capacity results with low redundant data.

The contributions of this work is twofold. Firstly, the realistic large-scale urban vehicular mobility trace characterization performed is important to help researchers on proposing novel dissemination protocols to vehicular networks. Moreover, the results showed which features are important in the dissemination selection process. The second contribution is the proposal of two heuristics that are a step forward to efficient, reliable and scalable dissemination schemes for vehicular networks. We hope these contributions will help on the development of new solutions to vehicular networks.

As future work, it is important to (i) implement and evaluate the proposed heuristics on simulation tools and in different mobility scenarios, and (ii) propose other heuristics based on network topology foresight.

REFERENCES

[1] T. Willke, P. Tientrakool, and N. Maxemchuk, "A survey of inter-vehicle communication protocols and their applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 2, pp. 3–20, 2009.

[2] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey," *IEEE Vehicular Technology Magazine*, vol. 2, pp. 12–22, June 2007.

[3] A. Bujari, "A Survey of Opportunistic Data Gathering and Dissemination Techniques," in *IEEE International Conference on Computer Communications and Networks (ICCCN)*, (Munich, Germany), pp. 1–6, July 2012.

[4] K. Rostamzadeh and S. Gopalakrishnan, "Analysis of emergency message dissemination in vehicular networks," in *IEEE Wireless Communications and Networking Conference*, (Cancun, Quintana Roo), pp. 575–580, Mar. 2011.

[5] P. Ruiz, B. Dorronsoro, P. Bouvry, and L. Tardón, "Information dissemination in VANETs based upon a tree topology," *Ad Hoc Networks*, vol. 10, pp. 111–127, Jan. 2012.

[6] Y. Sung and M. Lee, "Light-Weight Reliable Broadcast Message Delivery for Vehicular Ad-Hoc Networks," in *IEEE 75th Vehicular Technology Conference (VTC Spring)*, (Yokohama), pp. 1–6, May 2012.

[7] G. Ciccarese, M. De Blasi, P. Marra, C. Palazzo, and L. Patrono, "On the Use of Control Packets for Intelligent Flooding in VANETs," in *IEEE Wireless Communications and Networking Conference (WCNC)*, (Budapest), pp. 1–6, Apr. 2009.

[8] L. A. Villas, A. Boukerche, R. B. Araujo, A. A. F. Loureiro, and J. Ueyama, "Network partition-aware geographical data dissemination," in *EEE International Conference on Communications (ICC)*, pp. 1439–1443, 2013.

[9] W. Viriyasitavat, O. Tonguz, and F. Bai, "UV-CAST: an urban vehicular broadcast protocol," *IEEE Communications Magazine*, vol. 49, pp. 116–124, Nov. 2011.

[10] C. Xia, D. Liang, H. Wang, M. Luo, and W. Lv, "Characterization and modeling in large-scale urban DTNs," in *IEEE 37th Conference on Local Computer Networks (LCN)*, (Clearwater, FL), pp. 352–359, Oct. 2012.

[11] S. Ahmed and S. K. Salil, "Characterization of a large-scale delay tolerant network," in *IEEE 35th Conference on Local Computer Networks (LCN)*, (Denver, CO), pp. 56–63, Oct. 2010.

[12] R. Monteiro, S. Sargento, W. Viriyasitavat, and O. K. Tonguz, "Improving VANET protocols via network science," in *2012 IEEE Vehicular Networking Conference (VNC)*, pp. 17–24, IEEE, Nov. 2012.

[13] S. Uppoor and M. Fiore, "Insights on metropolitan-scale vehicular mobility from a networking perspective," in *4th ACM International Workshop on Hot Topics in Planet-scale Measurement*, (Low Wood Bay, Lake District, UK), pp. 39–44, 2012.

[14] G. Resta and P. Santi, "On the Fundamental Limits of Broadcasting in Wireless Mobile Networks," in *IEEE INFOCOM*, (San Diego, CA), pp. 1–5, Mar. 2010.

[15] M. J. Khabbaz, W. F. Fawaz, and C. M. Assi, "Which Vehicle To Select?," *IEEE Communications Letters*, vol. 16, pp. 812–815, June 2012.

[16] G. Zyba, G. M. Voelker, S. Ioannidis, and C. Diot, "Dissemination in opportunistic mobile ad-hoc networks: The power of the crowd," in *IEEE INFOCOM*, (Shanghai), pp. 1179–1187, Apr. 2011.

[17] S. Ahmed and S. Kanhere, "Hubcode: hub-based forwarding using network coding in delay tolerant networks," *Wireless Communications and Mobile Computing*, vol. 13, pp. 828–846, May 2011.

[18] S. Uppoor and M. Fiore, "Large-scale urban vehicular mobility for networking research," *2011 IEEE Vehicular Networking Conference (VNC)*, pp. 62–69, Nov. 2011.

[19] L. Cheng, B. Henty, D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 ghz dedicated short range communication (dsrc) frequency band," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1501–1516, 2007.

[20] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *7th ACM SIGCOMM Conference on Internet Measurement*, (New York, NY, USA), pp. 29–42, 2007.

[21] F. A. Silva, T. R. M. Silva, R. Vicente, L. B. Ruiz, and A. A. F. Loureiro, "On the improvement of vehicular macroscopic mobility models," in *IEEE Vehicular Technology Conference (VTC Fall)*, 2013.