Ad Hoc Networks xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Ad Hoc Networks

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

Geo-localized content availability in VANETs

Q1 Fabrício A. Silva [a,c,*], Azzedine Boukerche [b], Thais R.M. Braga Silva [c], Linnyer B. Ruiz [d], Antonio A.F. Loureiro [a]

[a] Department of Computer Science, Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil
[b] PARADISE Research Lab, University of Ottawa, Ottawa, Canada
[c] Department of Computer Science, Campus Florestal, Federal University of Viçosa (UFV), Florestal, MG, Brazil
[d] Department of Informatics, State University of Maringá (UEM), Maringá, PR, Brazil

A R T I C L E I N F O

Article history:
Received 24 November 2014
Revised 5 June 2015
Accepted 8 June 2015
Available online xxx

Keywords:
Vehicular ad-hoc networks
Content replication
Content availability

A B S T R A C T

Vehicular Ad Hoc Networks (VANETs) are emerging as a very useful tool for gathering, processing, and providing data to vehicles and passengers. It is expected that vehicles equipped with a variety of sensors will play a determining role in Intelligent Transportation System (ITS) and Smart City applications. With the evolution of VANET services comes the need for solutions to increase the availability of content to users. To this end, we propose a **G**eo-Localized **O**rigin–**D**estination-based **C**ontent **R**eplication (GO-DCR) solution that relies on vehicles' origin and destination points to decide which of them are more appropriate to replicate content inside a region of interest. We compare GO-DCR with two existing solutions through extensive simulations. The results show that GO-DCR increases content availability and reduces the cost of delivery.

1. Introduction

Vehicular Ad Hoc Networks (VANETs) are emerging as a very useful tool for gathering, processing, and providing data to vehicles and passengers [1–3]. Vehicles equipped with a variety of sensors will play a determining role in Intelligent Transportation System (ITS) and Smart City applications, since they will be able to collect accurate, location-aware information to feed new services. In addition, emerging services may take advantage of vehicle-to-vehicle (V2V) communication capabilities to reduce cost by alleviating network infrastructure, while maintaining a high quality of experience (QoE) for their users.

With the evolution of VANET services, from simple safety alert message exchanging to advanced systems capable of sensing, processing, and sharing content, comes the need for solutions to make content available to vehicles and passengers. Furthermore, many VANET applications are considered to be location-aware, meaning that content in such cases is of interest only to vehicles in specific locations, referred to as region of interest (RoI). We refer to such content as geo-localized. Given that vehicles move constantly, the source vehicle of geo-localized content may move outside the RoI, resulting in a loss of information; meanwhile, other vehicles in the ROI may have an interest in the information. Thus, a new challenge derives from the need of persisting content inside a RoI so that all interested vehicles will receive it.

Content replication [4–6], a technique frequently adopted for web content on the Internet to solve performance issues, is a concept potentially attractive for solving the geo-localized content delivery problem in VANETs. The objective of a content replication solution is to select appropriate vehicles to keep geo-localized content inside the RoI, while they deliver the content to the interested vehicles. By doing so, we

* Corresponding author at: Department of Computer Science, Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil. Tel.: +55 3135363300 . Q2

E-mail addresses: fabricio.asilva@dcc.ufmg.br, fabaguiarsilva@gmail.com (F.A. Silva), boukerch@site.uottawa.ca (A. Boukerche), thais.braga@ufv.br (T.R.M. Braga Silva), linnyer@gmail.com (L.B. Ruiz), loureiro@dcc.ufmg.br (A.A.F. Loureiro).

http://dx.doi.org/10.1016/j.adhoc.2015.06.004
1570-8705/© 2015 Elsevier B.V. All rights reserved.

can expect an increase in content availability in the RoI, but a reduced delivery cost, since the less expensive V2V communication may be preferable to the vehicle-to-infrastructure (V2I). In addition, we can expect to alleviate the underlying communication infrastructure. Thus, we claim that content replication is worth exploiting for geo-localized content delivery in VANET applications.

Nevertheless, because of the particular characteristics of VANET networks, replicating content is easier said than done. First, VANETs typically have a highly dynamic topology, leading to constant changes in the network topology as a result of intermittent contacts. In addition, constant changes in network density are expected, ranging from low to high, depending on the location and time. The continuous mobility of the vehicle suggests that several different connections will be required to complete a content delivery. Furthermore, content may only be valid for a short period of time (e.g., during a traffic jam). Finally, VANET applications are expected to be deployed in large cities with hundreds of thousands of vehicles. While progress has been made in the field of geo-localized routing [7–9], few studies have proposed efficient geo-localized content replication solutions for VANETs.

In this work, we propose and evaluate a **G**eo-Localized **O**rigin–**D**estination-based **C**ontent **R**eplication (GO-DCR) solution designed for VANET applications. GO-DCR relies on the vehicles' origin-destination (O–D) points (i.e., departure and arrival points) to select those that are more likely to be effective in keeping content inside the RoI. It should be noted that O–D information is easily obtained from navigation systems that are expected to be included in most vehicles in the near future. The main idea is that geo-localized content can be effectively replicated using easily obtained information (i.e., O–D points) and low-cost algorithms, in contrast to the majority of solutions found in the literature. To evaluate our proposal, we compare it with two existing solutions named Push-and-Track [10] and Linger [11], by running exhaustive simulations and measuring content availability and delivery cost.

The remainder of this work is organized as follows. In Section 2, we present the most relevant related studies discussed in the literature. In Section 3, we define the geo-localized content replication problem and describe our proposal. We then present the simulation setup and results in Section 4. Finally, in Section 5 we conclude our study and present some future work.

2. Related work

Geographic location is of great importance to the provision of services to on-board users in a VANET. Thus, several studies have focused on so-called *geocast* routing [7–9,12]. However, in addition to routing schemes, it is important to decide how to replicate and persist content inside a RoI so that all target vehicles will be covered in a low-cost way.

In general, geo-localized content replication solutions rely on distributed comparable indexes, allowing the vehicles themselves to decide the most appropriate replicas. Maihöfer et al. [13] proposed electing a vehicle as a content replica based on its expected travel time in the RoI. More than one vehicle may be elected to reduce the cost of the election process. A new election process is begun when the previously selected vehicles leave the RoI. Similarly, ARM [14] is a framework for electing content replicas as well. The election is based on the distance from the target central point, the angle between the vehicle direction and the target central point, vehicle speed, and target area. Only one vehicle is elected for each content, and a new election round is begun when the vehicle is no longer deemed appropriate to act as a replica. Jerbi et al. [15] proposed that vehicles be selected to form a virtual geo-localized backbone. The vehicles that compose the virtual backbone for a region (i.e., an intersecting area), are then responsible for disseminating the geo-localized information. One of the backbone vehicles is selected to perform local broadcasts as it reaches the center of the RoI. Finally, Linger [11,16] is a protocol used to transmit information inside a RoI. To this end, Linger proposes an index that is computed locally by vehicles based on the distance to the center point of the RoI, the angle relative to the RoI center point, and the vehicles' speed.

All of the aforementioned solutions rely on distributed algorithms to compute a comparable index that is used to select the replica vehicles. Usually, the index represents the vehicle's characteristics such as direction, distance to the RoI, speed, and trajectory. Nevertheless, this approach generates a high overhead due to the distributed negotiation, as well as the decision process for the initiation of the next selection round. In contrast, GO-DCR relies only on O-D information and on efficient centralized algorithms, with a broader view of the network.

Existing studies also propose solutions to deliver content from one fixed source station to a given destination. Thus, a vehicle is selected to be the content carrier from the source to the destination. This is the case of On-Time [17] and OVS-OBRM [18]. On-Time [17] is a routing protocol for bus transportation systems. Based on the scheduled stops of each bus, an algorithm tries to maximize the delivery probability within a given period, and selects the best carrier bus. Similarly, but considering all vehicles in a highway scenario, OVS-OBRM [18] selects a vehicle to be the carrier of content that must be delivered from one roadside unit (RSU) to another. Thus, OVS-OBRM proposes the computation of the residual travel time: the time taken for each vehicle to reach the target RSU. The vehicle in the vicinity of the source with the smallest residual travel time is then selected as the carrier.

Unlike GO-DCR, the main drawback of those solutions is the assumption that both source and destination are fixed entities. This is not the case in most VANET applications; to the contrary, the source or the destination move at a significantly high speed.

Similarly to GO-DCR, some studies propose infrastructure-based solutions that take advantage of powerful servers with a broader view of the network. In HomeZone [19,20], content is transmitted to a region by *Infostations* located there, as well as vehicles passing by or expected to pass by it. To keep the information present in the region of interest, vehicles are selected based on a utility function that evaluates a vehicle's ability to keep a replica in the RoI. This is done by a current content carrier that checks whether one of its neighbors may be a better carrier, and, in a positive case, transfers the content to it. The drawbacks of HomeZone is the overhead caused by periodic messages, and the need to know the vehicle's trajectory. The expected contact-graph is exploited in MobTorrent [21], in which infrastructure stations adopt graph algorithms with the objective of maximizing the transferred data. Likewise, the contact-graph is also a fundamental tool in [22] for selecting vehicles to carry content for delivery to other vehicles. In TEG-PW [23], mobility prediction is exploited to select replicas to prefetch content parts based on expected consumer-producer encounters. Those solutions rely on the knowledge of network topology, which is difficult to obtain in a highly dynamic network such as VANETs. In addition, the graph algorithms they adopt are usually expensive in terms of time and resources, particularly when many thousands of vehicles are providing as input. In contrast, our solution relies on O–D points (which are easily obtained from navigation systems), and less expensive algorithms.

Another study proposes Push-and-Track [10], which keeps track of vehicles that have already received content and decides whether to re-inject new replicas into the network. Vehicles in Push-and-Track send *Enter* messages when entering a RoI, *Leave* messages when leaving it, and *Ack* after being covered (i.e., received content). Thus, the server tracks how many vehicles are yet to be covered, then decides how many new replicas should be allocated. If this expectation is not satisfied, the server randomly selects new replicas and sends the content to them using the cellular network. Furthermore, Push-and-Track considers a panic zone period, in which the server sends the content to all uncovered vehicles inside the RoI through infrastructure communication. Since Push-and-Track relies on assumptions similar to GO-DCR, we adopt it as a baseline solution in our evaluation study.

All aforementioned studies have the same objective as ours: to persist content inside a RoI. However, many of them rely on distributed algorithms that cause network overhead, particularly in large-scale, highly dense VANETs. On the other hand, some solutions rely on infrastructure stations and on vehicle trajectory, a resource-consuming data to be obtained. Our proposal of GO-DCR relies on origin-destination points, easily obtained by navigation systems, to select the most appropriate replica vehicles to persist content inside a RoI.

3. Geo-localized content replication in VANETs

As mentioned earlier, many VANET applications are designed to deliver content to vehicles in a specific RoI. Hence, to improve content availability and the chances that all target vehicles (i.e., those traveling through the RoI within the content lifetime) will receive the content, it is important to replicate it on strategically selected vehicles. In this section, we formally describe this problem and propose a solution to it.

3.1. Problem statement

Let V and S be the sets of vehicles and infrastructure stations, respectively. The graph $G(V \cup S, E)$ represents the connectivity among vehicles themselves, as well as among vehicles and infrastructure stations. An edge $e_{i,j} \in E$ indicates that vehicle $v_i \in V$ has a contact with another vehicle (V2V) or an infrastructure station (V2I) $vs_j \in V \cup S$, where $v_i \neq vs_j$.

Also, let A be a circular area representing the RoI, with radius r and centered at point (p, q). The objective of a geo-localized content delivery application is to deliver content C to all vehicles $v \in V$ that travel through A during the content lifetime that starts at t_c^i and ends at t_c^f.

One possible option to cover all target vehicles is to send content to all vehicles entering the RoI through V2I communication. However, the cost of transmitting content using V2I communication is significantly higher than the cost of transmitting the same content using V2V communication. Thus, a content replication solution may be adopted to reduce the cost, but a high coverage rate must be maintained by selecting a subset $R \subseteq V$ of vehicles to act as localized content replicas, or in other words, surrogate servers.

In short, the objective of a geo-localized content replication solution is to select appropriate vehicles to help with the delivery of content to other vehicles that travel through a RoI. The replica vehicles should be selected so that content availability increases and delivery cost decreases. Below, we describe our proposal, GO-DCR, to solve this problem.

3.2. GO-DCR

We assume a hybrid scenario, where vehicles are equipped with cellular as well as WAVE (Wireless Access for Vehicular Environment) [24] communication modules. This is a reasonable assumption, since most vehicles are expected to be connected to the Internet in the near future [25]. This is a trend supported by both academia [26] and industry [27]. Therefore, the system model we assume will make the deployment of our solution quite simple. In this system model, vehicles are responsible for gathering data from their sensors and making it available to others. To help with this task, they take advantage of Internet servers with a broader view of the network. The server is responsible for selecting the vehicles to act as replicas.

Content is assumed to be static and is provided by a vehicle that has sensed geo-localized useful information, such as a traffic jam situation. This vehicle uses a cellular network to send the sensed content and its location reference to the server, which then proceeds to the execution of the replication task. At time of departure, ordinary vehicles send a message containing their O–D points, also using the cellular network. The assumption that vehicles know their destination points is reasonable given the increasing market penetration of online navigation systems, such as *Waze* [28], in which users provide their destination even for known routes to avoid traffic congestion. The adoption of such systems is expected to increase even more with the advance of solutions that provide real-time traffic status [29]. Furthermore, context-aware navigation systems [30] are expected to learn users' routine and infer their destination. In fact, *Waze* already does that and suggests a destination to the users depending on their current location and period of day.

Different approaches can be used to strategically select replicas. In this work, we use the origin o_i and destination d_i points of vehicle v_i to decide whether or not it is expected to be a good replica. Furthermore, our solution measures a coverage index that indicates how well the RoI is covered by replicas over time, in order to decrease delivery cost.

The GO-DCR content replication process runs every σ seconds in the server. For each execution, the server evaluates the efficiency of vehicles that sent their O-D points after the last execution round for their roles as replicas. A vehicle is

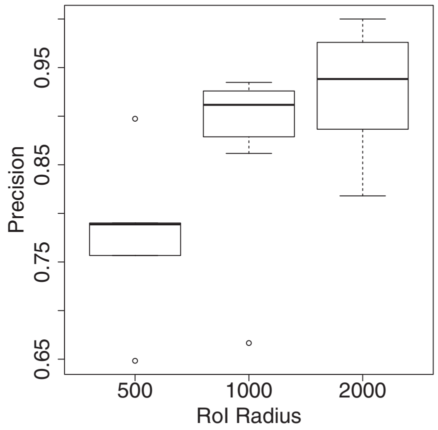


Fig. 1. Precision of vehicles that indeed travel through a RoI when a straight line segment from its origin to its destination intercepts the RoI.

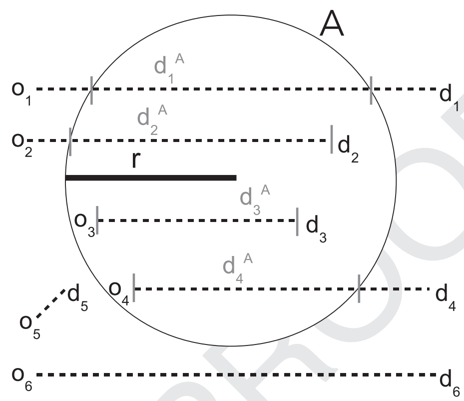


Fig. 2. Different scenarios of vehicles intercepting the area A based on a straight line from their O–D points.

only considered a replica candidate if it is expected to pass through the RoI, based on its O–D points.

A vehicle v_i is considered to travel through the RoI when a straight line segment from its origin o_i to its destination d_i intercepts the circular area A that represents the RoI. We assume vehicles travel following a straight-line segment from their origin to their destination to save resources and time. In this case, there is no need to compute the vehicle's route and trajectory, which requires complex graph algorithms. Although some vehicles may be incorrectly selected as replicas even when they do not travel through the RoI, we argue that this is acceptable due to conservation of time and resources. To demonstrate this rationale, we compute the ratio between the number of vehicles that travel through the RoI considering their real trajectory over the number of those that are assumed to do so in a straight line. In other words, To this end, we consider 33 randomly generated RoI of radius 500 m, 1000 m and 2000 m for the Cologne scenario [31,32] comprised of over 120,000 vehicles in a 400 km^2 area. In fact, as shown in the box plots of Fig. 1, the precision is quite good, particularly for larger RoIs. Therefore, only a small number of replica candidates, based on the straight line, fail to travel through the RoI in their real trajectories.

Two metrics contribute to the replica selection in GO-DCR:

- the distance d_i^A a vehicle v_i travels inside the RoI A: the longer the better, because vehicles that travel for longer distances inside the RoI are more prone to cover more vehicles;
- a coverage index I_i^A for the interval starting at time t_i^e when v_i enters A and ending at t_i^l when it leaves. This is inversely proportional to the number of vehicles covering the RoI in the interval $t_i^e - t_i^l$: the higher the better, because fewer vehicles will be simultaneously covering A.

Fig. 2 depicts examples of six possible scenarios of departure and arrival points. In the following, we refer to this figure whenever necessary in helping understand the solution.

To compute the distance that a vehicle travels inside A, we define the line formula that represents the segment $\overrightarrow{o_i d_i}$ as $y = mx + c$. Also, it is known that a circle A centered at point $(p,\ q)$ with radius r is defined as $(x^2 - p) + (y^2 - q) = r^2$. Next, we find the interception points of the line in the circle, if any. To this end, we first compute the values of m and c by replacing the o_i and d_i coordinates into the line formula. Next, we replace y from the line formula into the circle formula, obtaining $(x^2 - p) + ((mx + c)^2 - q) = r^2$. To conclude, we then solve the quadratic equation to find the points (x_i^1, y_i^1) and (x_i^2, y_i^2) where the line intercepts the circle.

In the event that no results are found, the vehicle will not pass through A, and is then not considered as a replica candidate, as seen in the example of Vehicle 6 in Fig. 2. Also, it should be noted that having interception points does not imply that the vehicle travels through A, since the line $y = mx + c$ is an extension of the segment $\overrightarrow{o_i d_i}$, as shown by Vehicle 5 in Fig. 2. Although the line representing the segment $\overrightarrow{o_5 d_5}$ intercepts A, the segment itself does not. We then compare the segment $\overrightarrow{o_i d_i}$ to the points where the line intercepts the circle to check whether or not $\overrightarrow{o_i d_i}$ is inside A partially or totally. The distance d_i^A is then computed as the Euclidian distance of points (x_i^1, y_i^1) and (x_i^2, y_i^2) which is the part of $\overrightarrow{o_i d_i}$ that is indeed inside A. After this procedure, we have d_i^A for Vehicle v_i, as marked for Vehicles 1–4, in Fig. 2. We then compute the contribution of this distance to the vehicle's selection as the percentage of this distance relative to A's diameter:

$$w_i^d = \frac{d_i^A}{2 \times r}. \tag{1}$$

The next step is to compute the coverage index I_i^A that measures how well A is covered during the period Vehicle v_i is expected to be inside it. To this end, let A_v be the area covered by each vehicle, which is given by the V2V communication range. We then estimate the number n_i^A of replica vehicles that will be inside A simultaneously with v_i, and the total coverage area that they are able to achieve in the best scenario (i.e., with no overlaps). Since the best scenario has a very small chance of happening, we expect the occurrence of redundant coverage areas. Then, we set the index as a negative quadratic function:

$$I_i^A(n_i^A) = \left(\frac{-1}{10}\right) \times \left(\frac{n_i^A * A_v}{A}\right)^2 + 100. \tag{2}$$

Please cite this article as: F.A. Silva et al., Geo-localized content availability in VANETs, Ad Hoc Networks (2015), http://dx.doi.org/10.1016/j.adhoc.2015.06.004

The reasoning is to have a higher chance of a vehicle being a replica when A is under cover. The idea of using a negative quadratic function is to be more flexible under a low number of simultaneous replicas, and more rigid in a selection already covered by numerous vehicles.

Finally, we compute the probability of a vehicle's selection as a replica by averaging both values defined by Eqs. (1) and (2):

$$p_i = \frac{w_i^d + I_i^A}{2}. \quad (3)$$

In summary, vehicles that travel for longer distances inside the RoI, when few other vehicles are expected to be there, are more prone to have higher selection probabilities. As soon as a selected replica enters the RoI, it starts disseminating content to its neighbors periodically, every δ s.

4. Performance evaluation

To assess the performance of GO-DCR as compared to two existing solutions, we conduct extensive simulations following two complementary approaches. The first, referred to as a large-scale study, adopts a large-scale mobility model and assumes an ideal network scenario with guaranteed packet delivery. To this end, we implemented a simulator in the *R* environment that performs efficiently with large-scale mobility data, by assuming an ideal network scenario in which packets are not lost or corrupted, and the transmission rate is always as high as configured. In addition to the large-scale evaluation, we also implement and compare both solutions in the OMNET++ network simulator, which is referred to as a network-enabled study.

We measure two major metrics: content availability and delivery cost. The former refers to the ease of content availability to interested vehicles, in terms of coverage, time to be covered, and capacity. The coverage is defined as N_c/N_t, where N_t is the number of target vehicles and N_c is the number of target vehicles covered. The time to be covered for vehicle v_i is computed as $TC_i - TD_i$, where TD_i is the time that v_i departs from its origin point, and TC_i is the time when it is covered. Finally, the capacity represents the amount of data that could be transmitted by replica vehicles, and is computed as $\Sigma_{i,j} d_{i,j} \times T_{V2V}$ $\forall i \in R, j \in V$, where $d_{i,j}$ is the duration of contacts between v_i and v_j, R is the set of replica vehicles, V is the set of all vehicles, and T_{V2V} is the transmission rate of the V2V communication technology.

The latter major metric, delivery cost, refers to the number of messages exchanged, and the amount of redundant that is transmitted. The number of V2I or infrastructure messages N_{V2I} is the number of all messages exchanged between infrastructure stations and vehicles (i.e., Enter, Leave, Ack, content from the server to selected replicas). The number of ad hoc or V2V messages N_{V2V} is the total number of messages exchanged between vehicles (i.e., control messages in Linger, content from replicas to vehicles). The redundant messages for a vehicle v_i is defined as $R_i = NC_i - 1$, where NC_i is the number of times v_i received the same content. The total of the redundant messages is then the sum for all vehicles, defined as $R_{total} = \sum_{i \in V} R_i$. In addition, we measure the number of network lost packets for the network-enabled study to assess how each solution affects network performance. In summary, the higher the coverage and the capacity, and the lower the time to be covered, the redundancy, and the cost in general, the better.

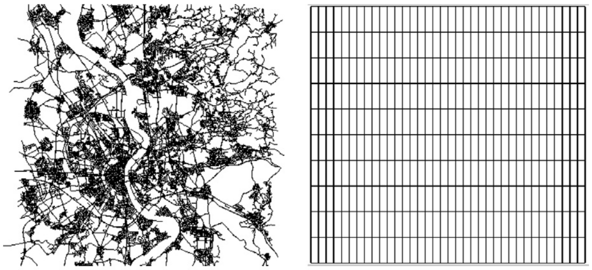
We present the simulation details and results for each study as follows. All results represent the mean and the 95% confidence interval of 33 simulation runs. For each run, a random RoI position is used, which is the same for all three solutions. This way, the solutions are compared under the same conditions, and for different scenarios. It is important to state that a different seed is used for the random number generator for each simulation run.

4.1. Baseline solutions

We compare GO-DCR to Push-and-Track [10] and Linger [11] through extensive simulation. In Push-and-Track, the infrastructure server keeps track of all target vehicles that are already covered (i.e., have received content). A vehicle sends the server an *Enter* message when it becomes a target for content, and a *Leave* message when it is no longer a target. By the time a vehicle receives the content, it sends an *Ack* message to the server. The server keeps track of the covered vehicles, and is then aware of the number of vehicles that are still uncovered. Furthermore, the server expects a linear coverage behavior, in which at least p% of the target vehicles are expected to be covered after p% of the content lifetime has elapsed. If this coverage expectation is not satisfied, the server randomly selects new replicas and sends them content using the cellular network. The replicas then start delivering the content periodically every δ seconds. Here, we consider δ to be 1 s, which is the same value used for GO-DCR. Push-and-Track also defines a panic zone that starts at a defined time before the content expiration time, when all uncovered vehicles receive the content through the cellular network. Here, the panic zone begins 10 s. before the expiry of content lifetime.

Linger [11], on the other hand, is a totally distributed solution in which vehicles compute a comparable index that indicates how suitable they seem to be as replicas. The index takes into account the vehicle's speed, direction, and distance to the RoI. The higher the index, the better replica a vehicle is expected to be. Initially, the vehicle that first senses or generates the geo-localized data to be shared is assumed as replica. From this time on, a distributed replica selection process starts with the first replica computing and sending its index to its one-hop distance neighbors. Upon receiving the index, these neighbors also compute their own indices, and compare them with the received value. When the computed index is higher than the received one, the replica candidate waits for a period that is inversely proportional to its index before responding to the current replica, indicating that it should become a replica. Thus, the first vehicle that responds is expected to have the higher index among the neighbors, and is then selected as the new replica. It should be noted that many replicas may exist simultaneously, and that each is responsible for looking for new, better replicas. Each replica delivers content to its neighbors every δ s. Here, we consider δ to be 1 s, which is the same value used for GO-DCR and Push-and-Track.

Please cite this article as: F.A. Silva et al., Geo-localized content availability in VANETs, Ad Hoc Networks (2015), http://dx.doi.org/10.1016/j.adhoc.2015.06.004



(a) Large-scale scenario (b) Network-enabled scenario

Fig. 3. Simulation scenarios.

4.2. Large-scale study

VANET solutions, particularly those regarding content replication, require large-scale performance evaluations to be proven effective; otherwise, the results may not represent reality. However, due to computational complexity and hardware constraints, the existing network simulators, such as NS-2 and OMNet++, do not perform well in large-scale mobility scenarios [33]. Thus, we developed a specific simulator in the *R* environment that performs efficiently with large-scale mobility data; this is achieved by assuming an ideal network scenario.

The simulations were performed for a publicly available, large-scale realistic mobility trace [31,32] from the city of Cologne, Germany, comprised of over 120,000 vehicles between 6:00 a.m. and 8:00 a.m. of a weekday, in a 400 km^2 area (See Fig. 3(a)). We vary the RoI radius from 500 m to 2000 m to measure how well the solutions perform under different application demands.

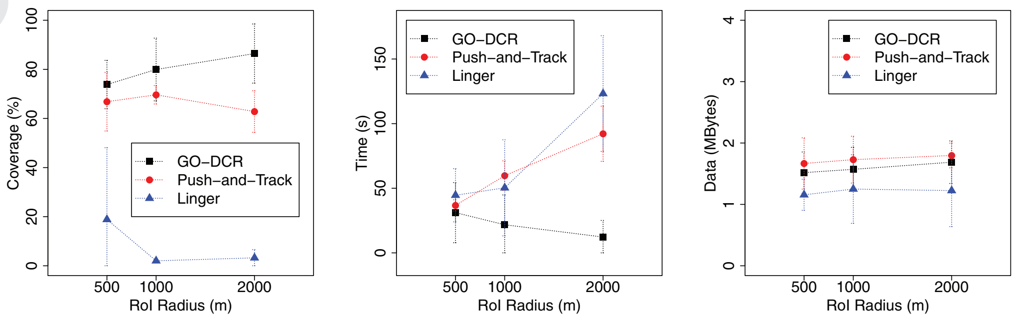
We assume V2V and V2I transmission rates to be of 1 Mbps and 3 Mbps, respectively, based on real experiments conducted in [34]. Vehicles have a transmission range of 100 m based on the results presented in [35]. Hence, $A_v = \pi \times 100^2$ is the area covered by each vehicle. We consider a static content of size 100 Kbytes, as a single fragment of the same size, to be delivered to vehicles that are inside the RoI during the 2 h of the mobility scenario. For all solutions, the replicas deliver the content periodically every $\delta = 1$ s when inside the RoI, and the server runs its allocation process every $\sigma = 1$ s. In Push-and-Track, the panic zone starts 10 s before the content lifetime expires. The number of simultaneous replicas assumed for Linger is 200 for the large-scale, and four for the network-enabled studies.

4.2.1. Results

Fig. 4 presents the content availability results. GO-DCR covered more vehicles than Push-and-Track and Linger, as shown in Fig. 4(a). This is because GO-DCR balances the number of replicas over time, and selects vehicles that are expected to be more valuable in terms of coverage. The coverage of Push-and-Track decreases for the 2000 m RoI radius scenario. In contrast, GO-DCR adapts accordingly to larger RoIs, because of the computed coverage index. It should be noted that Linger presents very poor coverage results in large-scale evaluations, particularly for larger RoI radius. Our network-enabled study confirms the proof provided by Linger's authors of its effectiveness for small mobility scenarios. However, Linger performs more poorly when large-scale mobility scenarios are used. In fact, given the large number of vehicles to be covered, Linger cannot select the appropriate replicas, covering only a small subset of all target vehicles. It is important to state that not even the panic zone strategy of Push-and-Track could improve its coverage; many vehicles leave the RoI without being covered, and then are not considered in this panic stage. Thus, this strategy is only effective for vehicles that are inside the RoI when the panic zone starts.

The appropriate replica allocation also led GO-DCR to cover vehicles earlier than Push-and-Track and Linger, as shown in Fig. 4(b). This metric is of great value when content must be delivered as soon as possible, as in the case for traffic condition alerts. Fig. 4(c) illustrates that Push-and-Track could transmit slightly larger content than GO-DCR and Linger, since contacts among replicas and target vehicles last longer. However, this accomplishment is undervalued due to the lower coverage obtained.

GO-DCR also performed better than Push-and-Track when it comes to delivery cost in terms of the number of V2I messages exchanged, as shown in Fig. 5(a). Given this result, we argue that the objective of giving preference to less expensive communication, V2V, was achieved by GO-DCR, as illustrated in Fig. 5(b). Since Linger does not rely on infrastructure stations, it does not require infrastructure messages



(a) Coverage (b) Time to be Covered (c) Capacity

Fig. 4. Large-scale simulation results in terms of content availability. GO-DCR achieves higher coverage and lower time to be covered results, when compared to Push-and-Track and Linger.

Please cite this article as: F.A. Silva et al., Geo-localized content availability in VANETs, Ad Hoc Networks (2015), http://dx.doi.org/10.1016/j.adhoc.2015.06.004

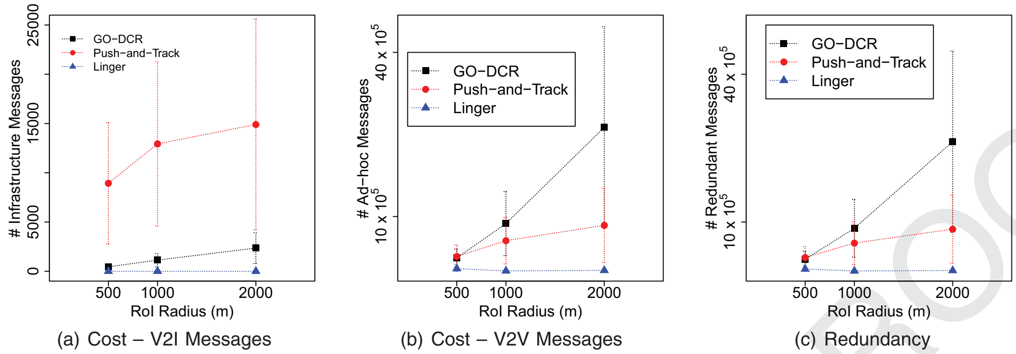


Fig. 5. Large-Scale simulation results in terms of communication cost. GO-DCR leads to a low number of V2I messages. However, the higher coverage achieved led to a higher redundant, in terms of V2V messages.

to operate. Linger also does not require many V2V messages, a result of the bad replica allocation.

However, the improvement in content availability and the preference for V2V communication come at a price. GO-DCR initiated a greater amount of redundant data than Push-and-Track and Linger because the replica vehicles were in more frequent contact with the same target vehicles, as shown in Fig. 5(c). Regardless, the redundant data in GO-DCR is a result of V2V content transmissions, which is less critical than using the most expensive V2I communication.

4.3. Network-enabled study

The objective of this study is to evaluate GO-DCR when vehicular specific network protocols are used. To this end, we implement the three solutions in the OMNET++[1] network simulator; it provides the WAVE (Wireless Access for Vehicular Environment) suite [24] which includes the IEEE 802.11p standard [36] for MAC and physical layers, and the IEEE 1609 protocol suite to define the upper-layer operations.

Due to computational constraints, we adopt a smaller mobility scenario from a different city to complement the large-scale results, then increase the evaluation reliability. Adopting a different mobility scenario from the large-scale study is important in order to assess the performance in two scenarios, instead of only one. We adopt a Manhattan-like mobility scenario, illustrated in Fig. 3(b), comprised of vertical and horizontal double-lane roads in a 9 km^2 area, in which blocks have an average size of 80 $\times$ 270 m^2. Manhattan was chosen because of its similarity with many other cities around the world in terms of roads and traffic. To simulate realistic vehicle movements, we take advantage of the mobility model defined by SUMO.[2] We fixed the radius of the RoI at 500 m to be in accordance with the simulated area. We consider a static content size of 100 Kbytes, as a single fragment of the same size, to be delivered to vehicles that are inside the RoI during the entire application running time.

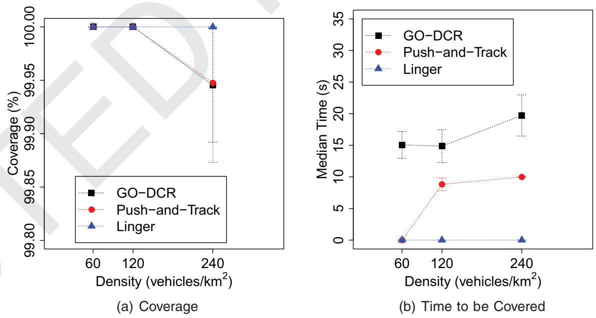


Fig. 6. Network-enabled simulation results in terms of content availability.

Based on the results presented in [37], we set the parameters of the physical layer as follows: the simplified path loss model exponent α parameter is 2, and the transmission power is 10 mW. To improve even more the realism of the physical layer, we also adopt the shadowing model described in [38]. This is a realistic model for urban environments based on IEEE 802.11p measurements that simulates signal attenuation caused by buildings. In other words, the communication is affected by the presence of buildings, as in real environments. We set the other parameters, such as delivery interval and content size, with the same values as in the large-scale evaluation study. We run simulations in low, medium, and high network density to evaluate different scenarios.

4.3.1. Results

The network-enabled content availability results are shown in Fig. 6. Given the confidence interval, all solutions achieved 100% coverage for all network densities, as shown in Fig. 6(a). This result was expected due to the size of the network scenario, which increases the chance of a target vehicle coming into contact with a replica vehicle. When it comes to the time-to-be-covered, Linger and Push-and-Track present slightly better values, particularly for low network densities, as depicted in Fig. 6(b). This is due to the fact that GO-DCR does not focus on covering vehicles as soon as possible, but on balancing the number of replicas along time and space. As a result, target vehicles at the beginning of the content

[1] http://www.omnetpp.org.
[2] http://sumo-sim.org.

Please cite this article as: F.A. Silva et al., Geo-localized content availability in VANETs, Ad Hoc Networks (2015), http://dx.doi.org/10.1016/j.adhoc.2015.06.004

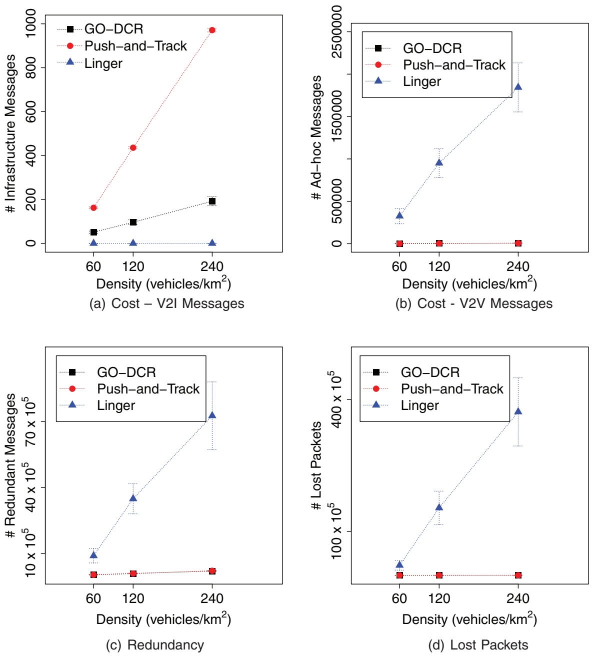


Fig. 7. Network-enabled simulation results in terms of cost (V2I messages, V2V messages, redundancy and network lost packets).

lifetime may not be in immediate contact with a replica vehicle. In addition, Push-and-Track selects more replicas than GO-DCR, which increases the chance of target vehicles making contact with replica vehicles. In fact, we measured that Push-and-Track selects on average approximately 3.7 times more replica vehicles than GO-DCR. Linger also performs quite well on this metric, regardless of the network density. For such a small-scale scenario, the distributed index-based approach followed by Linger is very useful. However, when we also consider the large-scale evaluation results, GO-DCR seems more attractive. Furthermore, these results come at a price for Push-and-Track, as shown by the communication cost results in Fig. 7.

With respect to communication costs, Fig. 7(a) shows that Push-and-Track requires a significantly higher number of V2I messages to be exchanged during its operation, when compared to GO-DCR and Linger. In fact, this is a key point, since V2I communication is more expensive than V2V. Again, Linger requires no V2I messages, since it relies exclusively on V2V communication.

On the other hand, both Push-and-Track and GO-DCR present similar results for V2V communication costs and redundancy, with slightly better values for GO-DRC, as shown in Fig. 7(b and c). These results are mainly because Push-and-Track selects more replica vehicles (i.e., approximately 3.7 times on average) and adopts the panic zone approach using V2I communication. On the other hand, GO-DCR balances the number of replicas over time to give more opportunities for using V2V communication. Linger, on the other hand, requires a larger number of V2V, and consequently, generates more redundant messages than the other solutions. Furthermore, these values increase in a manner linear to with the network density, since Linger requires more negotiation messages in the replica selection process.

Finally, Fig. 7(d) shows that Linger also results in greater packet loss, mainly due to communication congestion caused by the large number of messages required. This is even more critical for scenarios with greater network density.

It is important to discuss differences observed when comparing large-scale with network-enabled results. For example, Linger performed quite well in the small-scale, network-enabled scenario, and very poorly in the large-scale scenario. In other words, Linger, as originally proposed, should be used in small scenarios and avoided in large-scale ones. Therefore, we argue that each solution has its advantages and disadvantages.

In general, considering both large-scale and network-enabled results, we can state that our proposal, GO-DCR, achieves high coverage results and conserves resources by requiring a low number of V2I exchanged messages. We accomplished our objective of balancing the number of replicas over time, covering as many vehicles as possible and using as little infrastructure as possible. In conclusion, GO-DCR is a cost-effective solution that could be adopted in various scenarios.

5. Conclusion

Many VANET applications require content to be delivered to vehicles inside a specific RoI. To assist with this task, we propose a geo-localized content replication solution that selects appropriate vehicles to keep replicas and deliver them to the target vehicles. Our solution, called GO-DCR, relies on origin-destination points to allocate vehicles that are more likely to be good replicas. Extensive simulation results showed that, in general, GO-DCR performed better than two existing solutions in terms of content availability and delivery cost.

When it comes to content replication for VANET applications, there is still a great deal of work to be done. It will be important to evaluate GO-DCR considering different application demands, such as large and delay-sensitive content. In addition, incentive mechanisms, such as discounts in services and priority for data downloading, should be adopted to give benefits to vehicles that cooperate for the delivery process.

Acknowledgment

This work is supported by NSERC, Canada Research Chair Program and NSERC-DIVA Strategic Research Network, by CNPq under grant 573.738/2008-4 (INCT NAMITEC), and by CAPES under grant 99999.002061/2014-07.

References

[1] C. Zhang, R. Lu, X. Lin, P.-H. Ho, X. Shen, An efficient identity-based batch verification scheme for vehicular sensor networks, in: Proceedings of the IEEE INFOCOM, 2008, pp. 816–824, doi:10.1109/INFOCOM.2008.58.
[2] S.-C. Hu, Y.-C. Wang, C.-Y. Huang, Y.-C. Tseng, A vehicular wireless sensor network for co2 monitoring, in: Proceedings of the IEEE Sensors, 2009, pp. 1498–1501, doi:10.1109/ICSENS.2009.5398461.
[3] H. Conceição, M. Ferreira, J. Barros, On the urban connectivity of vehicular sensor networks, in: Distributed Computing in Sensor Systems, Springer, 2008, pp. 112–125.

[4] G. Peng, CDN: content distribution network, 2004, arXiv preprint cs/0411069.

[5] A. Vakali, G. Pallis, Content delivery networks: Status and trends, Internet Computing, IEEE 7 (6) (2003) 68–74.

[6] A. Pathan, R. Buyya, A Taxonomy and Survey of Content Delivery Networks, Technical Report, Grid Computing and Distributed Systems Laboratory, Melbourne, 2007.

[7] S. Allal, S. Boudjit, Geocast routing protocols for vanets: Survey and geometry-driven scheme proposal, J. Internet Serv. Inf. Secur. (JISIS) 3 (1/2) (2013) 20–36.

[8] A. Bujari, A Survey of Opportunistic Data Gathering and Dissemination Techniques, in: Proceedings of the International Conference on Computer Communications and Networks (ICCCN), IEEE, Munich, Germany, 2012, pp. 1–6, doi:10.1109/ICCCN.2012.6289225.

[9] C. Rezende, A. Mammeri, A. Boukerche, A.A. Loureiro, A receiver-based video dissemination solution for vehicular networks with content transmissions decoupled from relay node selection, Ad Hoc Netw. 17 (2014) 1–17. http://dx.doi.org/10.1016/j.adhoc.2013.12.011.

[10] J. Whitbeck, Y. Lopez, J. Leguay, V. Conan, M.D. de Amorim, Push-and-track: saving infrastructure bandwidth through opportunistic forwarding, Pervasive Mob. Comput. 8 (5) (2012) 682–697, doi:10.1016/j.pmcj.2012.02.001.

[11] M. Fiore, C. Casetti, C.-f. Chiasserini, D. Borsetti, Persistent localized broadcasting in VANETs, IEEE J. Sel. Areas Commun. 31 (9) (2013) 480–490, doi:10.1109/JSAC.2013.SUP.0513043.

[12] A. Boukerche, Algorithms and protocols for wireless, mobile ad hoc networks, 77, John Wiley & Sons, 2008.

[13] C. Maihöfer, T. Leinmüller, E. Schoch, Abiding geocast: time–stable geocast for ad hoc networks, in: Proceedings of the International Workshop on Vehicular Ad Hoc Networks, ACM, 2005, pp. 20–29, doi:10.1145/1080754.1080758.

[14] D. Borsetti, M. Fiore, C. Casetti, C.-F. Chiasserini, An application-level framework for information dissemination and collection in vehicular networks, Perform. Eval. 68 (9) (2011) 876–896, doi:10.1016/j.peva.2010.11.004.

[15] M. Jerbi, A.-L. Beylot, S.M. Senouci, Y. Ghamri-Doudane, Geo-localized virtual infrastructure for VANETs: design and analysis, in: Proceedings of the Conference on Global Telecommunications, IEEE, 2008, pp. 1–6, doi:10.1109/GLOCOM.2008.ECP.131.

[16] D. Borsetti, M. Fiore, C. Casetti, C.F. Chiasserini, When mobile services go local, in: Proceedings of the International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), ACM, Tenerife, Canary Islands, Spain, 2009, pp. 235–244, doi:10.1145/1641804.1641845.

[17] U. Acer, P. Giaccone, D. Hay, G. Neglia, S. Tarapiah, Timely data delivery in a realistic bus network, in: Proceedings of the International Conference on Computer Communications (INFOCOM 2011), IEEE, Shanghai, China, 2011, pp. 446–450, doi:10.1109/INFCOM.2011.5935201.

[18] M.J. Khabbaz, W.F. Fawaz, C.M. Assi, Which vehicle to select? IEEE Commun. Lett. 16 (6) (2012) 812–815, doi:10.1109/LCOMM.2012.042512.112350.

[19] I. Leontiadis, Publish/subscribe notification middleware for vehicular networks, in: Proceedings of the Symposium on Middleware Doctoral, ACM, 2007, pp. 1–6, doi:10.1145/1377934.1377936.

[20] I. Leontiadis, P. Costa, C. Mascolo, A hybrid approach for content-based publish/subscribe in vehicular networks, Pervasive Mob. Computing 5 (6) (2009) 697–713, doi:10.1016/j.pmcj.2009.07.016.

[21] B.B. Chen, M.C. Chan, MobTorrent: a framework for mobile internet access from vehicles, in: Proceedings of the IEEE Conference on Computer Communications, IEEE, 2009, pp. 1404–1412, doi:10.1109/INFCOM.2009.5062056.

[22] O. Trullols-Cruces, M. Fiore, J. Barcelo-Ordinas, Cooperative download in vehicular environments, IEEE Trans. Mob. Computing 11 (4) (2012) 663–678, doi:10.1109/TMC.2011.100.

[23] F. Malandrino, C. Casetti, C.-F. Chiasserini, M. Fiore, Content download in vehicular networks in presence of noisy mobility prediction, IEEE Trans. Mob. Computing (99) (2013) 1, doi:10.1109/TMC.2013.128.

[24] Y.L. Morgan, Notes on dsrc & wave standards suite: Its architecture, design, and characteristics, Commun. Surv. Tutor. IEEE 12 (4) (2010) 504–518.

[25] N. Lu, N. Cheng, N. Zhang, X. Shen, J.W. Mark, Connected vehicles: solutions and challenges, IEEE Internet Things J. 1 (4) (2014) 289–299.

[26] J. Barros, How to build vehicular networks in the real world, in: Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing, ACM, 2014, pp. 123–124.

[27] Telefonica, Connected car industry report, (2014).

[28] Waze, 2015, (Online (available at http://www.waze.com)).

[29] M.B. Younes, A. Boukerche, G. Rom'an-Alonso, An intelligent path recommendation protocol (icod) for vanets, Comput. Netw. 64 (2014) 225–242.

[30] A. Ramazani, H. Vahdat-Nejad, A new context-aware approach to traffic congestion estimation, in: Proceedings of the 4th International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, 2014, pp. 504–508.

[31] S. Uppoor, M. Fiore, Insights on metropolitan-scale vehicular mobility from a networking perspective, in: Proceedings of the International Workshop on Hot Topics in Planet-scale Measurement, ACM, Low Wood Bay, Lake District, UK, 2012, pp. 39–44.

[32] S. Uppoor, M. Fiore, Large-scale urban vehicular mobility for networking research, in: Proceedings of the 2011 IEEE Conference on Vehicular Networking (VNC), IEEE, 2011, pp. 62–69, doi:10.1109/VNC.2011.6117125.

[33] S. Joerer, F. Dressler, C. Sommer, Comparing apples and oranges? in: Proceedings of the ACM International Workshop on Vehicular Inter-Networking, Systems, and Applications, ACM Press, New York, USA, 2012, pp. 27–32, doi:10.1145/2307888.2307895.

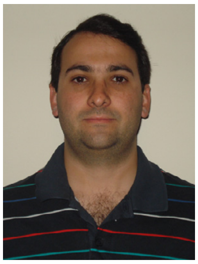
[34] F.A. Teixeira, V.F. e Silva, J.L. Leoni, D.F. Macedo, J.M. Nogueira, Vehicular networks using the ieee 802.11 p standard: an experimental analysis, Veh. Commun. 1 (2) (2014) 91–96.

[35] L. Cheng, B. Henty, D. Stancil, F. Bai, 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 J. Sel. Areas Commun. 25 (8) (2007) 1501–1516.

[36] D. Jiang, L. Delgrossi, Ieee 802.11p: towards an international standard for wireless access in vehicular environments, in: Proceedings of the IEEE Conference on Vehicular Technology, 2008, VTC Spring, 2008, pp. 2036–2040, doi:10.1109/VETECS.2008.458.

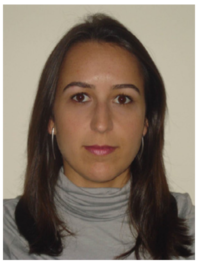
[37] C. Sommer, S. Joerer, F. Dressler, On the applicability of two-ray path loss models for vehicular network simulation, VNC 12 (2012) 64–69.

[38] C. Sommer, D. Eckhoff, R. German, F. Dressler, A Computationally Inexpensive Empirical Model of IEEE 802.11p Radio Shadowing in Urban Environments, Technical Report CS-2010-06, University of Erlangen, Dept. of Computer Science, 2010.



Fabrício A. Silva received his B.S. degin Computer Science from Federal University of Minas Gerais, Brazil, in 2004. He received his M.Sc. degree in 2006 and is currently a Ph.D. Candidate in the Computer Science Department at the Federal University of Minas Gerais. In 2014, he held a visiting researcher position at PARADISE Research Laboratory, University of Ottawa. His research interests are in the areas of mobile and ad-hoc networks, specially vehicular ad hoc networks.

Azzedine Boukerche is a Full Professor and holds a Canada Research Chair position in distributed simulation and wireless and mobile networking at the University of Ottawa. He is the Founding Director of PARADISE Research Laboratory at Ottawa U. Prior to this, he held a faculty position at the University of North Texas, USA. He worked as a Senior Scientist at the Simulation Sciences Division, Metron Corporation located in San Diego. He was also employed as a Faculty at the School of Computer Science McGill University, and taught at Polytechnic of Montreal. He spent a year at the JPL/NASA-California Institute of Technology where he contributed to a project centered about the specification and verification of the software used to control interplanetary spacecraft operated by JPL/NASA Laboratory.



Thais R.M. Braga Silva holds a B.Sc., M.Sc., and Ph.D. in Computer Science from Federal University of Minas Gerais (UFMG), Brazil. Currently she is a Professor of Computer Science at Federal University of Viçosa, Brazil. Her research interests include ubiquitous and mobile networks, with emphasis on context-aware computing.

Linnyer B. Ruiz is a Computer Science Professor at State University of Maringá (UEM), Brazil. She holds a Ph.D. in Computer Science received from Federal University of Minas Gerais, Brazil (2003) and a M.Sc. degree in Electrical Engineering and Industrial Informatics received from Technological University of Paranaá (1996). Her main research efforts are devoted to various aspects of embedded systems, wireless sensor networks, ubiquitous computing, ubiquitous healthcare, and vehicular networks. She has published several papers in the area of wireless ad hoc and sensor networks.

Antonio A.F. Loureiro received his B.Sc. and M.Sc. degrees in computer science from the Federal University of Minas Gerais (UFMG), Brazil, and the Ph.D. degree in computer science from the University of British Columbia, Canada. Currently, he is a full professor of computer science at UFMG, where he leads the research group in mobile ad hoc networks. His main research areas are wireless sensor networks, mobile computing, and distributed algorithms. In the last 10 years he has published regularly in international conferences and journals related to those areas, and also presented tutorials at international conferences.

Please cite this article as: F.A. Silva et al., Geo-localized content availability in VANETs, Ad Hoc Networks (2015), http://dx.doi.org/10.1016/j.adhoc.2015.06.004