Performance Modeling of Vehicular Floating Content in Urban Settings

Gaetano Manzo
University of Bern & HES-SO Valais
Switzerland
gaetano.manzo@hevs.ch

Marco Ajmone Marsan
Institute IMDEA Networks, Spain &
Politecnico di Torino, Italy
marco.ajmone@imdea.org

Gianluca A. Rizzo
HES-SO Valais
Switzerland
gianluca.rizzo@hevs.ch

Abstract—Among the proposed opportunistic content sharing services, Floating Content (FC) is of special interest for the vehicular environment, not only for cellular traffic offloading, but also as a natural communication paradigm for location-based context-aware vehicular applications. Existing results on the performance of vehicular FC have focused on content persistence, without addressing the key issues of the effectiveness with which content is replicated and made available, and of what are the conditions which enable acceptable FC performance in the vehicular environment. This work presents a first analytical model of FC performance in vehicular networks in urban settings. It is based on a variation of the random waypoint (RWP) mobility model, and it does not require a model of road grid geometry for its parametrization. We validate our model extensively, through numerical simulations on real-world traces, showing its accuracy on a variety of mobility patterns and traffic conditions. Through analysis and simulations, we show the feasibility of the FC paradigm in realistic urban settings over a wide range of traffic conditions.

I. INTRODUCTION

As a consequence of the progressive realization of the IoT vision, recent years have witnessed an explosion of data generated at the edge of the network, raising scalability issues in current computing and communication approaches, that cannot be addressed efficiently by traditional centralized paradigms. Applications with tight delay constraints, such as mobile offloading, or autonomous coordinated driving, require content to be stored and processed as close as possible to the end user, hence at the network edge.

For those applications, typically, content is mainly of use when in the proximity of data sources. This push towards the edge of the network is also at the origin of the interest in opportunistic communications and ad-hoc networking, which offload the communication infrastructure through direct peer-to-peer exchanges of data.

An example of opportunistic communication scheme for the local dissemination of information goes under the name of Floating Content (FC) [1] or Hovering Information [2]. It enables probabilistic content storing in geographically constrained locations - denoted as *Anchor Zones* (AZ) - and over a limited amount of time. Given the infrastructure-less nature of the paradigm, and its reliance on opportunistic exchanges among mobile nodes, a significant portion of the performance studies on FC has focused on the conditions under which content persists in the AZ (i.e., "floats" for a "large enough"

amount of time). Paper [1] characterizes the *critical condition*, for the content to float indefinitely with very high probability, under various mobility models. [3] introduces an analytical model for content persistence for the case of outdoor pedestrian mobility in large open spaces, such as a city square. However, for any practical application, content persistence over time within the AZ is only a necessary condition for FC viability. Indeed, when the content persists, it is essential to characterize how often the FC paradigm manages to deliver the content to the intended users (i.e., how effectively the content stored probabilistically can be retrieved). Despite its importance, few works consider this issue. For setups where nodes move according to a random direction mobility model, [4], [5] characterize analytically the *success probability* (i.e., the probability of delivering content to users in transit in the AZ) as a function of system parameters. [5] shows however that those results are heavily inaccurate when applied to realistic scenarios, and in particular to vehicular setups. As we already noted, FC is of special interest for the vehicular environment, not only as a way to offload cellular networks traffic on Vehicular Ad-hoc Networks (VANETs), but also as a natural communication paradigm for location-based context-aware vehicular applications, such as traffic and accidents warnings, in which the relevance of the information is strongly correlated with proximity in space.

Indeed, several works consider FC in the vehicular environment [6], [7]. However, these works still focus on content persistence within the AZ, without addressing the key issue of what are the conditions under which FC is feasible in the vehicular environment. Hence, how the features of vehicular mobility patterns affect FC performance, and how to engineer a vehicular application relying on FC for achieving a target performance, are issues which remain to date still open.

In this work, we present a first step towards addressing these issues. We focus on urban scenarios, and we propose an analytic approach to performance evaluation of FC in vehicular scenarios in a city. Our approach is based on mapping the patterns of vehicular mobility in an urban environment into a modified version of the random waypoint (RWP) mobility model. This allows the derivation of a performance model which is not based on any specific road grid geometry. Numerical assessments of our results on measurement-based mobility traces show that our approach accurately models

vehicular FC performance over a wide range of values of system parameters. Through simulations on real-world vehicular traces, we perform a first characterization of the relationship between the main activity of city districts (office/commercial, industrial, or residential), their induced vehicular mobility patterns, and FC performance. As expected, we observe that the nature of mobility plays a crucial role in determining the achievable success probability. Contrary to what was shown in the literature, results from simulations suggest that in realistic urban vehicular settings the critical condition is not a good indicator of the probability for the content to float over time. The paper is organized as follows. In Section II we present the system model, and in Section III we present the main analytic results on FC success probability. In Section IV, we numerically assess the accuracy of our analytic results, and we evaluate them by simulation on a realistic setup. In Section V we briefly discuss previous work, and finally, Section VI concludes the paper.

II. SYSTEM MODEL

We consider a set of wireless nodes moving on a region of the plane, with transmission range r. We assume two nodes come in contact when they are in range of each other, i.e. when their distance is not larger than r (Gilbert's model [8]).

In what follows, we focus on an urban environment. Coherently with typical features of vehicles mobility in an urban setting, we assume that mobility patterns of each node alternate between time intervals spent moving, and time intervals spent still at a waypoint, which could model a pause at a crossroad, or in a parking lot.

More specifically, we assume nodes move according to the *District mobility model*, which abstracts from the details of the geometry of the street grid. Its formal definition is as follows:

Definition 1 (District mobility model). *We assume that nodes arrive in a region of the plane according to a Poisson process with intensity λ. Nodes, inside the region, move according to the RWP mobility model with pause, with velocity v constant and equal for all nodes (though our approach can be easily generalized to scenarios where velocity varies between waypoints according to a given distribution).*
When a node arrives, its initial waypoint is chosen uniformly at random on the border of the region. Upon reaching a waypoint, each node pauses for a random duration. Then, with probability p, the location of the next waypoint is selected uniformly at random within the region. Otherwise, with probability $1 - p$ the location of the next waypoint is selected uniformly at random on the border of the region. Once reached the border, a node disappears from the region.

As in RWP, in the district mobility model the sojourn time of every node within the region consists in a succession of *epochs*, where each epoch is the sum of the time spent moving between two consecutive waypoints, (the *moving time*), and of the time spent still at the destination waypoint (the *pause time*). The duration of a moving time, T_m and pause time, T_s, are assumed to be independent random variables. We denote

their pdf with f_{T_m} and f_{T_s}, respectively. We assume such pdfs to be the same for all nodes. One of the distinctive features of vehicular mobility is its being strongly influenced by the geometry of the road grids, and related constraints on node speed and direction. As a consequence, existing approaches to vehicular FC modeling focus on specific geometries, such as highways, highway junctions, or Manhattan grids, deriving results which are hardly generalizable [6].

Different from the typical vehicle mobility within a street grid, the district mobility model is isotropic (i.e., in any point in space, all directions can be chosen with equal probability) and unconstrained (i.e., a vehicle can occupy any location within the area). However, when a "sufficiently large" portion of the road grid is included inside an AZ (say, a few blocks), even if the possible instantaneous directions of movement are finite, for any two points chosen at random on the map (on the road grid), often there exists at least one path between them (albeit usually not on a straight line). Hence, on macroscopic scale, and for the purpose of modeling content replication and diffusion dynamics within the AZ, these differences only impact the mean moving time between two waypoints.

A. Floating Content basic operation

We assume that at a time t_0, a node in the plane (the *seeder*) defines a circular area of radius R, the *Anchor Zone* (AZ), containing the node itself. At time t_0 the seeder (blue node in Fig. 1a) generates a piece of content (e.g., a text message, a picture). For $t \geq t_0$, every time a node with the content comes in contact with a node without it within the AZ, the content is exchanged successfully with probability Q (Fig. 1b). We assume the time taken to replicate the content is negligible with respect to contact time. Nodes entering the AZ do not possess a copy of the content, and those exiting the AZ discard their copy. As a result of such opportunistic exchange, the content "floats" (i.e., it persists probabilistically in the AZ even after the seeder has left the AZ). In this way, the content is made available to nodes traversing the AZ (Fig. 1c) for the whole duration of its floating lifetime.

A first performance parameter of FC is content *availability* at a given time, i.e. the ratio between the number of nodes with content over the total amount of nodes inside the AZ at that time. Indeed, a high value of availability is correlated with low likelihood of content disappearance from the AZ, but also with the probability of getting the content for a node entering the AZ. This last feature is captured by the *success probability*, which is the probability for a node entering the AZ of getting out of the AZ with a copy of the content. This is the main performance indicator, and the one which is most directly related to the performance of applications and services relying on FC. Indeed, it is a direct indicator of the effectiveness with which the floating content is delivered and made available to nodes traversing the area.

Among other relevant performance parameters, the mean time to get the content is also important, for safety applications which require that the message reaches the destination as soon as possible.
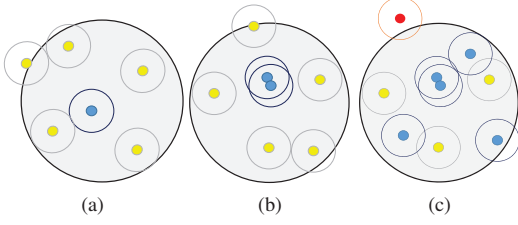
Fig. 1: Basic operation of Floating Content. 1a) Seeder (blue) defines the AZ. 1b) Opportunistic message exchange between nodes. 1c) Nodes going out of the AZ (red) discard the content.

III. AN ANALYTICAL MODEL FOR SUCCESS PROBABILITY

In this section, we present the derivation of the result which relates the success probability to the main system parameters. For our derivation, we consider a system in equilibrium, in which the mean number of nodes inside the AZ does not change over time. In what follows we assume, without loss of generality, a circular shaped district area with radius R. For nodes mobility within the AZ, we consider a district mobility model, with area coinciding with the floating content anchor zone. With q we indicate the mean fraction of the duration of an epoch in which a node moves:

$$q = \frac{E[T_m]}{E[T_m] + E[T_s]}$$

To derive the analytical expression for success probability, we start with the following two results, that determine the frequency with which any pair of nodes comes within range of each other in the AZ, and the content availability (and hence the mean number of nodes with content in the AZ).

Lemma 1 (Frequency of contacts). *In the district mobility model considering a region with area A, the mean frequency of contacts between any two nodes, with mean speed v and transmission range radius r, is*

$$\nu = \frac{2rqv(2(1-q) + 1.27q)}{A} \quad (1)$$

For the proof of Lemma 1 we refer to Appendix A.

The sojourn time for a node in the AZ is given by the number of epochs spent within the AZ with a probability p, plus the time to enter and exit the AZ with a probability $(1-p)$.

$$T_{soj} = (1-p)\sum_{k=0}^{\infty} p^k(kE[T_{epoch}] + E[T_m]) \quad (2)$$

Hence $\overline{N} = \lambda T_{soj}$ is the mean number of nodes in the AZ, by Little's law.

In general, the number of nodes with content in the AZ at a given time $t \geq 0$, $n(t)$, and without content, $m(t)$, are random variables. In what follows we indicate with $\bar{n}(t)$, and $\bar{m}(t)$ their stochastic average, and we consider what happens to these quantities when the transient due to the initial spreading

of the content within the AZ is concluded. Let $\bar{n}$, and $\bar{m}$ hence denote the values taken by the previous quantities *in stationary regime*. We have the following result.

Lemma 2 (Content availability, non-spatial model). *For $R >> r$, when the critical condition*

$$\overline{N}T_{soj}\nu > 1 \quad (3)$$

is satisfied, the stationary mean number of nodes in AZ with content, $\bar{n}$, is given by

$$\bar{n} = \bar{N} - \frac{1}{T_{soj}\nu Q}$$

The mean number of nodes without content, denoted with $\bar{m}$, is

$$\bar{m} = \frac{1}{T_{soj}\nu Q} \quad (4)$$

For the proof of Lemma 2 we refer to Appendix B. We now present our main result, based on the computation of the probability for a generic node to get the content during an epoch of its sojourn time inside the AZ, P_{epoch}.

Theorem 1 (Success Probability). *In the district mobility model, when $R >> r$ and the criticality condition (3) is satisfied, the probability for a node to get the content during its sojourn time within the AZ, in stationary regime is*

$$P_{succ} = \frac{P_{epoch}}{1 - p(1 - P_{epoch})} \quad (5)$$

where P_{epoch} is the probability that a node gets the content during an epoch (other than the final one), given by

$$P_{epoch} = P_m + (1 - P_m)P_s \quad (6)$$

P_s is the probability of getting the content during the pause time, given by

$$P_s = \int_0^{+\infty} \left(1 - e^{-\nu\tau\bar{n}Q}\right) f_{T_s}(\tau)d\tau \quad (7)$$

and P_m is the probability of getting the content during the moving time, given by

$$P_m = \int_0^{\frac{2R}{v}} \left(1 - e^{-\nu\tau\bar{n}Q}\right) f_{T_m}(\tau)d\tau \quad . \quad (8)$$

Where the pdf of T_m is given by

$$f_{T_m}(\tau) = \frac{4\tau v^2}{\pi R^2}\left(\arccos\frac{\tau v}{2R} - \frac{\tau v}{2R}\sqrt{1 - \left(\frac{\tau v}{2R}\right)^2}\right) \quad (9)$$

For the proof of Theorem 1 we refer to Appendix C.

Note that the epoch in which the node enters and exits the AZ coincides with the time spent moving towards the border of the AZ, as the node is assumed to disappear once reached the border. Hence for the final epoch $P_{epoch} = P_m$.

IV. SIMULATION AND VALIDATION

In this section, we present the numerical assessment of the FC behavior in a vehicular environment, as well as the validation of our model's accuracy, considering two scenarios. In a first scenario, we consider nodes moving according to the district mobility model, and we compare simulation results with the output of our model. Our goal is to characterize the FC performance as a function of the main system parameters, and to validate Theorem 1 against the system model, by assessing the impact of those effects, such as clustering of nodes with content and border effects, which are not included in our model.

In a second scenario, in order to perform a more realistic assessment, we consider a setup where mobile users are vehicles in the streets of a large European city, moving according to a set of measurement-based mobility traces.

In both scenarios, network simulations are performed using the VEINS simulation framework [9], based on OMNET++ [10] for network simulation, and on SUMO [11] for road traffic simulation. We assume that content is transferred (with probability $Q = 1$) every time two nodes come in range of each other, regardless of the amount of time spent in range. This corresponds to the case in which the content size is small enough for content transfer time to be much smaller than the typical mean contact time in urban scenarios. This is the case, for instance, of such applications as traffic warning, accident warning, or advertisements. We assume T_s to be exponentially distributed, with mean $1/\mu$.

Finally, we group vehicular services relying on FC into two broad categories. The "near real time" category typically floats messages with a very short validity in time (a few minutes). Examples are situated introductions, or infrastructure-less ridesharing, where passing cars inform neighboring pedestrians of their availability and their planned trip. Services in the "delay tolerant" category, instead, are associated to events with slower dynamics, and hence they require longer floating lifetimes of contents inside the AZ (of the order of one-two hours).

A. Baseline Scenario

In order to assess the accuracy of our analytic results, we performed a steady-state analysis of the system. Each simulation run has started with an empty AZ. In each simulation run, after waiting long enough for the transient on node population to be exhausted, we have assumed that the closest node to the center of the AZ generates a message, and starts replicating it opportunistically. After waiting long enough for the transient on content population in the AZ to be exhausted, we have measured content availability within the AZ, averaged over two hours. Being this a steady-state analysis, we only retained data for those contents which float. Moreover, for the sake of homogeneity in measured data, we retained only simulation results for which the content floats for the whole simulation time. Note that, since the critical condition was satisfied in all simulation setups, the vast majority of the content fluctuated for the whole simulation time.
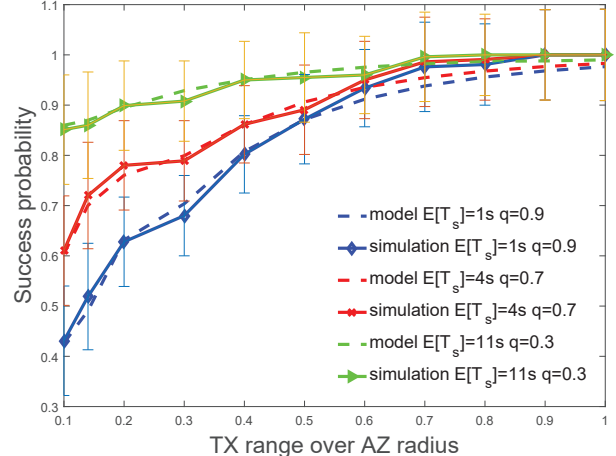

Fig. 2: Success probability vs transmission radius, with 95% confidence intervals.

In addition, for each simulation we have measured the average *success rate* over the simulation time. It is defined as the fraction of nodes which leave the AZ with a copy of the content during the simulation time, and it is an estimator of the success probability for the given content.

Unless otherwise stated, we considered an AZ radius $R = 50m$, arrival rate $\lambda = 0.1s^{-1}$, and a mean moving time $9s$. The probability p for a node to remain in the AZ after a pause time has been set to 0.9. In Fig. 2 and 3, respectively, we plot the values of success rate and availability derived from simulations, as a function of the ratio between the transmission range and the AZ radius, for different values of mean pause time $1/\mu$, and hence of the fraction of moving nodes q. In both plots, we compare simulation results with the respective values derived from our analysis.

As the plots show, both performance parameters increase monotonically as a function of r/R. Indeed, both are directly related to contact rate, and from Lemma 1 we know that contact rate is directly proportional to r/R.

The range of values of r/R was chosen as follows. Values of r/R below 0.1 do not allow the content to float for more than a few minutes (the critical condition is not satisfied). When $r/R > 1$ instead, even at very low node densities, nodes within an AZ form a strongly connected component, bringing success probability very close to one.

Indeed, when r/R gets closer to one, node clusters start to emerge, lowering the estimates of success rate from our model with respect to its actual value. However, this effect is mitigated by border effects. Indeed, in our derivation we have discarded the fact that users close to the border of the AZ have part of their coverage area lying outside of the AZ. Ignoring this effect brings to overestimate the replication opportunities of a content with the AZ. However such effect, typically negligible for low values of r/R, starts to have some impact on the accuracy of our model for values of r/R close to one.
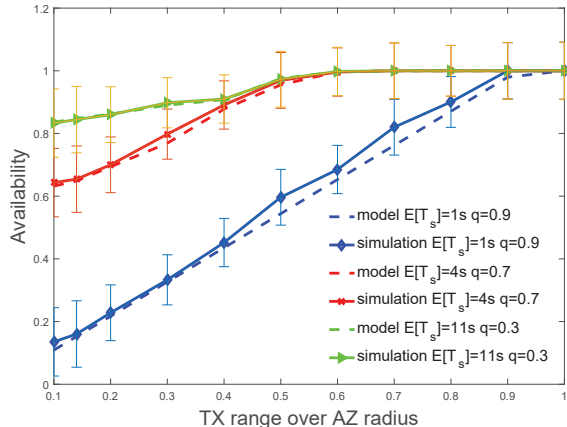
Fig. 3: Availability vs transmission radius, with 95% confidence intervals.

These plots show the good accuracy of our model for FC performance across various system configurations. Indeed the overall difference between analysis and simulation is always less than 5%. Note that, for the same value of r/R, keeping constant all the other system parameters, increasing $1/\mu$ has two contrasting effects on success probability. On the one side, it increases the mean sojourn time, increasing the mean number of nodes in the AZ, hence, at least in principle, the replication opportunities for the content. On the other side, it decreases the fraction q of nodes which at any time instant are moving within the AZ, thus decreasing the overall contact rate in the AZ. As we can see from the figure, the net result is an increase of success probability and success rate.

We also see that the marginal benefit of increasing r/R is higher for high values of q, indicating that the more mobile nodes there are, the higher are the advantages of a large transmission range.

Fig. 3 shows that availability has a similar dependency with respect to r/R and q as success probability. The figure also shows that there is not a direct proportionality between availability and success probability, and that typically availability is lower than success probability. Hence, as the plots suggest, and despite being often chosen as the main FC performance parameter, availability is, in general, a poor indicator of success probability. We can also see that, increasing the mean amount of time spent moving by a node allows decreasing the availability required for achieving a given value of success probability.

B. Luxemburg City simulations

In order to test our model in a realistic context, and to characterize those conditions in which FC is feasible in a vehicular environment, we considered a second scenario. It consists of an area of $155.95\ km^2$, which includes the city of Luxembourg and its surroundings. The street grid and the measurement-based mobility traces for this scenario for a 24-hour period were derived from [12]. Given the heterogeneity
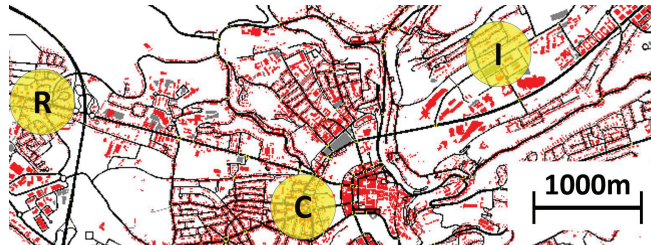

Fig. 4: Map of Luxembourg [12]. The yellow areas correspond to three AZ with $R = 260m$, located in the city center (C), in the industrial district (I), and in a residential district (R).

of the urban environment, in order to capture the effects on FC performance of a specific road structure, and of the resulting mobility patterns, we focused on three different locations as centers of AZs. One has been set in the city center (downtown, area "C" in Fig. 4). The second one in a residential area (area "R"), and the third in an industrial district (area "I"). Moreover, we considered two different time intervals over the course of the 24 hours. A first one, from 7AM to 9AM, corresponds to a period of peak traffic in the city, due to people commuting to work. A second time interval was chosen early in the morning, from 2.30AM to 5.30AM. Being this typically a period of very low car density in the whole city, it represents a worst case scenario for FC performance. For both time intervals, we chose a duration (two hours) which is of the same order of magnitude of the typical validity of messages in services such as traffic warnings, or accident warnings.

In Table I we list some of the main parameters of mobility for the three AZs, for an AZ radius of $260m$. These values show that the three districts differ substantially in terms of mean node density, mean sojourn time, and mean contact rates.

More specifically, the industrial district and the city center have a sensibly higher node density and contact rate than the residential district. Indeed, as the map in Fig. 4 shows, and different from the other two areas, the residential zone is mainly a transit zone for most vehicles, with little local traffic, and with few main roads carrying the majority of the traffic of the area. This implies shorter sojourn times and hence fewer opportunities for content replication.

Despite these differences, however, we observed that mean pause time, mean moving time, and mean speed are essentially the same for the three districts (and equal to about 15 s, 25 s, and $v = 14$ m/s, respectively), and they do not vary significantly in the course of the 24 hours. Note that, for the computation of the mean pause time, we assumed that cars which are parked do not participate in content exchange.

In each of the three districts and each simulation run, among all nodes passing in the AZ in the first 5 minutes of the given time interval, we chose at random a node as seeder. Then, from the moment in which the seeder enters the AZ, we simulated the process of content replication and floating for the whole duration of the time interval.

Figure 5 shows success probability for the three districts, for

TABLE I: Main mobility parameters for the three areas in Fig. 4. AZ radius $R = 260m$, transmission range $r = 20m$.

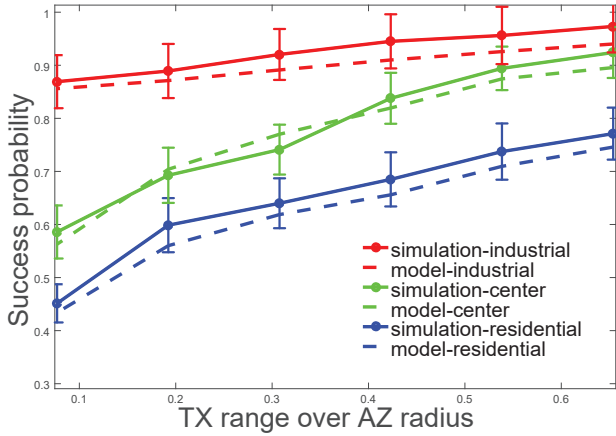| Zone		T_{soj} [s]	$\nu(N-1)$ [s^{-1}]	N	λ [s^{-1}]
C	low	335.7	0.04	29.5	0.087
	peak	368.4	1.16	489.2	1.327
I	low	412.2	0.05	38.7	0.0938
	peak	440.8	0.72	512.1	1.161
R	low	189.6	0.0063	1.9	0.010
	peak	192.4	0.043	12.9	0.067

Fig. 5: Success Probability vs transmission radius, in several locations of Luxembourg city, with 95% confidence interval.

an AZ radius of $260m$, and for the low traffic time interval, as a function of the ratio between transmission range and AZ radius. These plots show that, even in scenarios with realistic mobility patterns, simulation results are in good agreement with analytical values of success probability from Section III. The fact that our results maintain a high accuracy on very different road grid scenarios and mobility patterns suggests that the district mobility model is effective in capturing those aspects of vehicular mobility which have a significant impact on FC performance in urban settings.

In the industrial area, the node density and mobility is such that even with a small transmission range ($26m$) success probability is very high (above 0.8). As for the city center AZ, despite having only marginally lower node density, it exhibits a markedly smaller success probability at low values of r/R. Similarly to what seen in the baseline scenario, in scenarios with smaller sojourn times, increasing the ratio between transmission range and AZ radius brings to larger marginal increases in success probability. Finally, the plot indicates the residential area as the most critical area for FC performance, requiring a large ratio r/R to achieve high success probabilities. Note that in practical settings, values of r/R close to one defeat the purpose of having an AZ, as the vast majority of users who should get the content are within transmission range.

In order to get insight on FC performance in the Luxembourg scenario, for each of the three city districts, in Fig. 6 we plot (in red) the mean content availability over time for 30 simulation runs.

Content availability, being a property of the system in steady state, is not defined when content does not float. In experimental scenarios, the content is ultimately subject to disappearance, possibly due to stochastic fluctuations in the populations of nodes, or to specific mobility patterns, which bring out of the AZ those nodes with content, or which prevent content from being exchanged. In Fig. 6, we plot (in gray) the fraction of nodes with content over simulation time, for each of the 30 simulation runs. As we can see, in some configurations (e.g., in the city center AZ, for $R = 260m$ in the low traffic time interval), content does not float for the whole duration of the interval. For this reason, we have also plotted (in black in Fig. 6) the mean fraction of nodes possessing a copy of the content at a given time from content generation. This last quantity, taking into account content which disappears, is a better measure of the effective likelihood for a node inside the AZ to possess a copy of the content at a given time from content generation.

In all configurations, the critical condition (3) is satisfied. However, in the low traffic period with $R = 260$, only in the industrial district the content floats (in most of the cases) for the whole duration of the interval (Fig. 6a). In the city center AZ, instead, (Fig. 6e), despite achieving rapidly a high availability (0.7 within the first ten minutes), on average content floats for 30 minutes, while after an hour content has almost invariably disappeared. Similar behavior is observed in the residential area, where the content floats for at most 10 minutes (not shown).

Summarizing, in low traffic periods, even with relatively large ratios r/R, content floats for few minutes at most, making FC suitable only for "near real time" applications, with short lifetime. In the residential zone, at all times of the day, a large AZ radius is required to float the content with high probability for more than 30 minutes. In low traffic, extending the AZ radius while keeping constant the transmission radius has an overall positive impact on availability, and on floating lifetime, in all parts of the city (Fig. 6b, 6f and 6h). The critical condition is a poor indicator of actual feasibility of FC in realistic scenarios. When content floats for the whole two hours, we note a strong correlation between the evolution of availability of different content over time (Fig. 6a, 6b, 6c and 6f). In addition, we see availability taking only a finite set of values over time. This is an indication of the existence of few clusters of vehicles, with little inter-cluster exchanges. Only the increase of the number of vehicles arriving in the AZ (at around $4000s$ from beginning of floating time, in the low traffic time interval, see Fig. 7) increases the exchanges between these clusters (sharp increase in Fig. 6a, 6b and 6f). Finally, by appropriately setting the AZ radius (within reasonable limits, i.e. not extending it to the whole city) it is always possible to set content to float for a consistent amount of time.
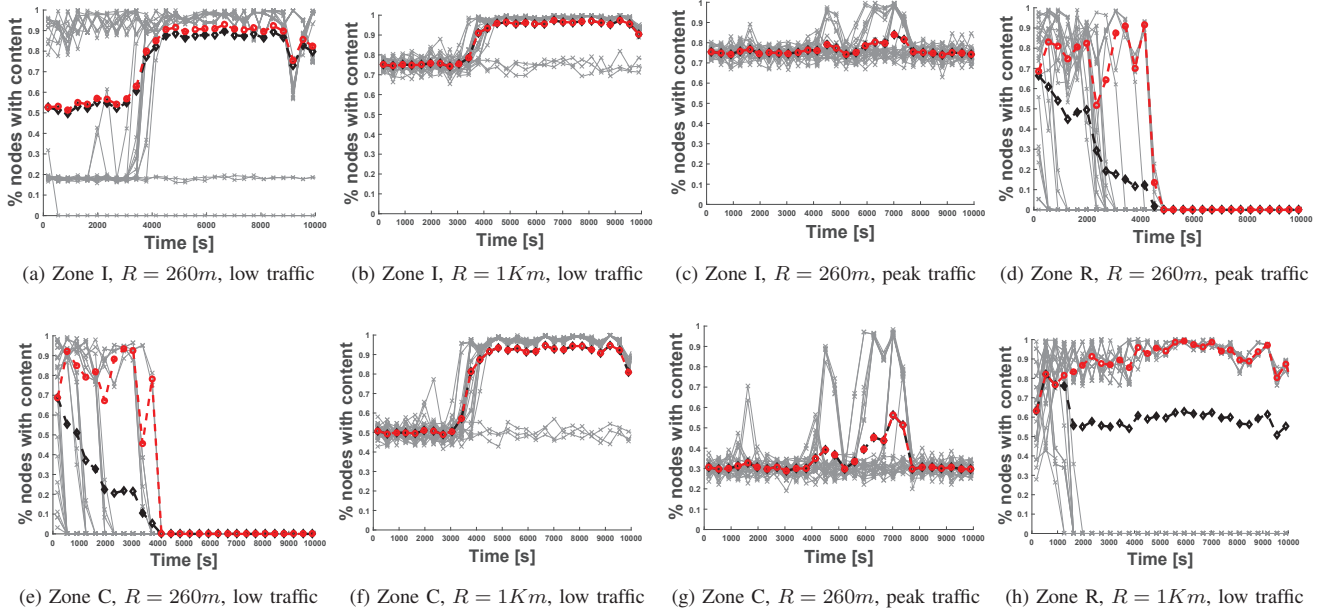
Fig. 6: Fraction of nodes with content (–x), mean fraction of nodes with content (–♦), and mean availability (–o) over floating time for the three city districts in Fig. 4. The low traffic time interval is 2.30 to 5.30AM, while peak traffic time interval is 7 to 9AM. Transmission radius is $100m$.
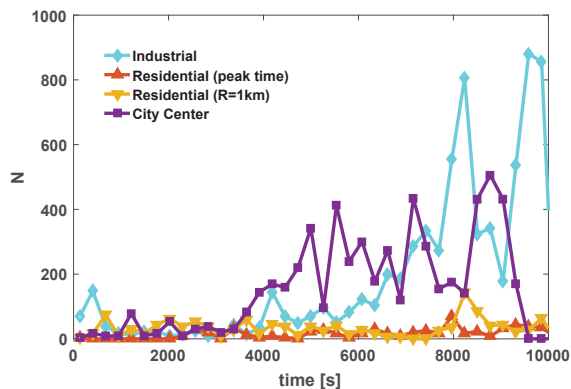

Fig. 7: Number of nodes in the three locations over simulation time. Industrial, city center and residential with R=1km are measured during low traffic time (2:30, 5:30). Residential peak traffic time (7:00, 9:00) has R=260m.

V. RELATED WORK

In [13] the authors consider a campus setup, and propose an analytic model of FC based on a Poisson jumps mobility model, which captures exchanges within user clusters, and between clusters. However, the derivation of the model is based on the assumption that on-the-fly exchanges between nodes on the move are negligible, in a way that results are not easily generalizable to other contexts, such as vehicular ones, where such assumption does not hold. For FC performance in vehicular environments, [6] develops an analytic model

for the mean floating lifetime on a two-lanes highway. [7] proposes a strategy for minimizing content replication within the AZ, while ensuring that the mean number of users with content within the region never gets below a given target value. Differently from these works, the analytic approach proposed in the present paper focuses on those aspects of FC performance, such as success probability, which are more closely related to the performance of applications which rely on FC.

VI. CONCLUSIONS

In this work, we presented a first analytical model for FC performance in vehicular networks, which does not require detailed modeling of the road grid geometry, but only a few key parameters of mobility. We assessed numerically our analytical results both on synthetic mobility patterns and on real-world vehicular traces, showing that they are in good agreement with simulation results. Through simulations, we have shown the feasibility of the FC paradigm in realistic urban setups on a variety of traffic conditions and mobility patterns. Our work gives a first indication of how to engineer a vehicular application based on FC, and on which vehicular applications are most suitable for FC in a given urban environment and time of the day.

In the followup to this work, we plan to investigate aspects such as the role of clustering in FC diffusion dynamics, of the ways in which pedestrians and infrastructure could be integrated into the FC paradigm, and how to extend it to non-urban scenarios (e.g. highways).

REFERENCES

[1] E. Hyytiä, J. Virtamo, P. Lassila, J. Kangasharju, and J. Ott, "When does content float? Characterizing availability of anchored information in opportunistic content sharing," in *INFOCOM*, Shanghai, China, Apr. 2011, pp. 3123–3131.

[2] A. A. V. Castro, G. D. M. Serugendo, and D. Konstantas, "Hovering information–self-organizing information that finds its own storage," in *Autonomic Communication*. Springer, 2009, pp. 111–145.

[3] M. Desta, E. Hyytia, J. Ott, and J. Kangasharju, "Characterizing content sharing properties for mobile users in open city squares," in *WONS 2013*, April 2013.

[4] S. Ali, G. Rizzo, B. Rengarajan, and M. A. Marsan, "A simple approximate analysis of floating content for context-aware applications," in *2013 Proceedings IEEE INFOCOM*, April 2013, pp. 21–22.

[5] S. Ali, G. Rizzo, M. A. Marsan, and V. Mancuso, "Impact of mobility on the performance of context-aware applications using floating content," in *Context-Aware Systems and Applications*, ser. Lecture Notes of the Institute for Computer Sciences. Springer International Publishing, 2014, vol. 128, pp. 198–208.

[6] K. Nakano and K. Miyakita, "Information floating on a road with different traffic volumes between opposite lanes," *Journal of Advanced Simulation in Science and Engineering*, vol. 3, no. 1, pp. 97–113, 2016.

[7] E. Koukoumidis, L.-S. Peh, and M. Martonosi, "Regres: Adaptively maintaining a target density of regional services in opportunistic vehicular networks," in *Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on*. IEEE, 2011, pp. 120–127.

[8] E. N. Gilbert, "Random Plane Networks," *Journal of the Society for Industrial and Applied Mathematics*, vol. 9, no. 4, pp. 533–543, 1961.

[9] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved ivc analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, 2011.

[10] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Simutools '08*. ICST, 2008, pp. 60:1–60:10.

[11] D. Krajzewicz, G. Hertkorn, C. Rössel, and P. Wagner, "SUMO (Simulation of Urban MObility), an open-source traffic simulation," in *Proceedings of the 4th middle East Symposium on Simulation and Modelling (MESM20002)*, 2002, pp. 183–187.

[12] L. Codeca, R. Frank, and T. Engel, "Luxembourg SUMO traffic (LuST) scenario: 24 hours of mobility for vehicular networking research," in *2015 IEEE Vehicular Networking Conference (VNC)*, Dec 2015, pp. 1–8.

[13] S. Ali, G. Rizzo, V. Mancuso, and M. A. Marsan, "Persistence and availability of floating content in a campus environment," *IEEE Conference on Computer Communication (INFOCOM)*, pp. 2326–2334, Mar. 2015.

[14] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Performance analysis of mobility-assisted routing," *in MobiHoc*, pp. 49–60, Jun. 2006.

[15] G. F. Newell, "The M/G/∞ queue," *SIAM Journal on Applied Mathematics*, vol. 14, no. 1, pp. 86–88, 1966.

[16] B. Christian, H. Hannes, and P. C. Xavier, "Stochastic properties of the random waypoint mobility model: Epoch time, direction, distribution and cell change rate," *ACM/Kluwer Wireless Networks, Special Issue on Modeling and Analysis of Mobile Networks*, vol. 10, no. 5, pp. 555–567, March 2003.

APPENDIX

A. Frequency of contacts

Proof. Let us consider two nodes a and b within the region, and the case in which $p = 1$, i.e. nodes never exit the region. Let us assume that node a moves according to the district mobility model, while node b is static. We consider the location of b as well as the initial location of a to be random within the region.

Let us divide a time of an epoch of node a into intervals of same duration τ, with τ much smaller than the expected value of T_m and of T_s, and small enough to assume that, in each interval, a is either moving or static. The mean fraction of intervals in which a moves is hence q. q is also the probability that a is moving during an interval of duration τ. Let us

compute now the mean number of contacts which node a experiments in a time interval $[0, \tau]$ during which a is moving. Note that since a moves along a straight line during a moving time, it cannot meet b more than once. Hence, the expected number of contacts between the two nodes is equal to 1 which multiplies the probability of finding b in the surface swept by a during the given time interval. As the location of b is uniformly distributed within the region, such probability is the ratio between the area of the surface swept by a, and the total area A.

Note that we assume an event of contact occurs at the first time instant in which two nodes are in range of each other. Hence in the computation of the contacts occurred in $[0, \tau]$, if node b is already in range of a at $t = 0$ we do not consider it as a contact occurred in the given time interval. As a consequence, for the computation of the above probability we have to subtract from the total area swept by a in the time interval, the area covered by a at $t = 0$. The resulting expression of the probability p_c that the two nodes come in contact during $[0, \tau]$, when a moves, is

$$p_c = \frac{2rv\tau}{A} \qquad (10)$$

In every interval, a moves with probability q. Hence, the mean number of contacts during $[0, \tau]$ is $p_c q$, and the contact rate is $p_c q / \tau$. Finally, the mean time between contacts is the inverse of the mean contact rate.

Let's now consider the case in which a and b both move without ever exiting the considered area. In each interval τ, the two nodes both move with probability q^2. For computing the mean time between contacts, we consider the equivalent setup in which a moves at a relative speed v_r while b is still. Due to the uniform choice of waypoint at every epoch, the expected value of v_r is given by

$$E_{v_r} = \frac{v}{2\pi} \int_0^{2\pi} \sqrt{(1 + \cos\theta)^2 + \sin^2(\theta)} d\theta \qquad (11)$$

which is equal to $1.27v$. With probability $2(1-q)q$ only one of the two nodes is moving (with speed v), and both do not move with probability $(1-q)^2$. Therefore, the mean node speed during a generic time interval is $1.27vq^2 + 2(1-q)qv$, and the mean contact time between the two nodes is

$$\frac{A}{2rqv(1.27q + 2(1-q))}$$

The inverse of this quantity is the frequency of contacts between any pair of nodes in our system, and we denote it with ν. If there are N nodes in the area, the overall contact rate is hence $\nu N(N-1)/2$.

When $p < 1$, in equilibrium the mean number of nodes does not vary. That means, when our system is in equilibrium state, ν is also the mean contact rate per couple of nodes. □

B. Node density

Proof. We write the balance equations for the mean number of nodes with (and without) content. The mean number of nodes with content is the result of nodes with content going

out from AZ, and of nodes without content coming in contact with nodes with content, and successfully exchanging it. If T_{soj} is the mean sojourn time in the AZ, and if we assume that nodes with content are uniformly distributed within the AZ, $n(t)/T_{soj}$ is the rate at which, at time t, nodes with content decide to get out of the AZ. The rate at which nodes acquire content is determined by the frequency with which a node with content comes in contact with a node without it, within the AZ. The mean rate of contacts within the AZ is given by ν (the mean rate at which a given pair of nodes comes in contact) multiplied by the number of node pairs within the AZ at t, $\frac{N(t)(N(t)-1)}{2}$. Of these contacts, those which increase $n(t)$ are those in which only one node of the two has the content, and for which the content is transferred successfully. As $p_n(t) = \frac{n(t)}{N(t)}$ the fraction of nodes with content in the AZ, and as Q is the probability that a content transfer is successful, the mean rate at which $n(t)$ increases at t is given by

$$2p_n(t)(1-p_n(t))Q\frac{N(t)(N(t)-1)}{2}\nu \qquad (12)$$

Since $N(t) \gg 1$, we can approximate (12) with $\nu n(t)m(t)Q$. Hence we have

$$\nu\bar{n}\bar{m}Q - \frac{\bar{n}}{T_{soj}} = 0 \qquad (13)$$

Over time, m increases due to arrivals of new nodes into the AZ, and it decreases because of nodes without content leaving the AZ, and as nodes without content come in range of nodes with content. Hence

$$\lambda - \nu\bar{n}\bar{m}Q - \frac{\bar{m}}{T_{soj}} = 0 \qquad (14)$$

By Little's law and we have balance equations we get (4). □

C. Success probability

Proof. We start by analyzing the probability for one or more successful content exchanges during a single epoch in the sojourn time of a node in the AZ. In the stationary regime, for the given mobility model, we assume that content is uniformly distributed among the population of nodes. This implies that at a given time instant the probability for a node to have the content is the same for all nodes, it is independent of node position within the AZ, and of the relative position of nodes. Such an assumption holds when clusters of more than two nodes are rare, which is typically the case when node densities are not high and $R \gg r$ [14].

We consider an epoch in the sojourn time of a node, and we evaluate the probability for a node to acquire the content during the moving time of that epoch. At time t, let $N(t)$ be the number of nodes in the AZ. Hence, there are $N(t)-1$ node pairs of which a given node is part. The mean contact rate at t is $\nu(N(t)-1) \approx \nu N(t)$. It can be easily seen that the system can be modeled as a $M/G/\infty$ queue, whose distribution of number of customers in the queue is Poisson [15]. Hence, in stationary state, $N(t)$ is Poisson with intensity $\overline{N}$. Therefore the expected contact rate for a single node at t (where the expectation is with respect to the contact rate between a pair of nodes) is also Poisson distributed, with intensity $\nu\overline{N}$. Note

that this holds because the two processes (contact rate for a pair of nodes, and number of nodes in the AZ) are independent. Thanks to the uniform content distribution, the probability for a node to have the content is $\frac{\bar{n}}{N}$. The probability for a node of having an unsuccessful contact (i.e., to get in contact with a node without content, or to fail in transferring a piece of content) is therefore $\frac{\bar{m}}{N} + \frac{\bar{n}}{N}(1-Q) = 1 - \frac{\bar{n}}{N}Q$. The probability for a moving node to get the content during T_m, conditioned to having j contacts is the complement of the probability that j contacts bring to no transfer of content. That is,

$$\mathbb{P}[succ, m \mid T_m, j] = 1 - \left(1 - \frac{\bar{n}}{N}Q\right)^j \qquad (15)$$

By the law of total probability,

$$\mathbb{P}[succ, m \mid T_m] = \sum_{j=1}^{\infty}\mathbb{P}[succ \mid T_m, j]Poiss_{\overline{N}\nu T_m}(j)$$

$$= 1 - \exp\left(-\overline{N}\nu T_m Q\frac{\bar{n}}{N}\right)$$

For the derivation of the pdf of the moving time f_{T_m}, note that T_m has a different distribution when it refers to the first moving time of a node in an AZ, when it refers to the last moving time in the AZ, and when the moving time is the first and the last of the node path inside the AZ (i.e., the node does not stop inside the AZ), and in all the other cases. However, when the mean number of waypoints on the path of a node is large enough, we can assume all these distribution to be the same as in those cases in which the node is neither entering nor exiting the AZ. The distribution of the transition length L given by [16]

$$f_L(l) = \frac{4l}{\pi R^2}\left(\arccos\frac{l}{2R} - \frac{l}{2R}\sqrt{1 - \left(\frac{l}{2R}\right)^2}\right) \quad .$$

Then, given that $T_m = \frac{L}{v}$, $f_{T_{m,d}}(\tau) = vf_L(\tau v)$. By the law of total probability, the probability of success during a moving time is hence

$$P_m = \int_0^{\frac{2R}{v}}\left(1 - e^{-\nu\tau\bar{n}Q}\right)f_{T_m}(\tau)d\tau \quad ,$$

In a similar manner, the probability for a static node to get the content during T_s is

$$P_s = \int_{\tau=0}^{+\infty}\left(1 - e^{-\nu\tau\bar{n}Q}\right)f_{T_s}(\tau)d\tau \quad . \qquad (16)$$

Finally,

$$P_{epoch} = P_m + (1 - P_m)P_s \qquad (17)$$

The probability of getting the content during the sojourn time in the AZ is given by

$$P_{succ} = (1-p)\sum_{k=1}^{\infty}p^{k-1}[1 - (1 - P_{epoch})^{k-1}] \quad , \qquad (18)$$

□