Contents lists available at ScienceDirect

Ad Hoc Networks

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

Stop and forward: Opportunistic local information sharing under walking mobility

Gianluca A. Rizzo^a, Vincenzo Mancuso^b, Shahzad Ali^c, Marco Ajmone Marsan^{b,d,*}

^a HES-SO Valais, Switzerland ^b IMDEA Networks Institute, Spain ^c Jouf University, Saudi Arabia ^d Politecnico di Torino, Italy

ARTICLE INFO

Article history: Received 5 March 2018 Accepted 21 May 2018 Available online 28 May 2018

Keywords: Floating content Analytical model Experiment Campus environment

ABSTRACT

In this paper we study an opportunistic geographically constrained information sharing paradigm known under the name Floating Content (FC), considering two different mobility models that describe the behavior of pedestrians. We assume that users carrying their smartphones walk from one location to another and then stop for a while. Information transfers take place in the periods when users pause between movements. We develop analytical models to compute the performance metrics that characterize FC in this case and we validate analytical results with data collected during an experiment performed in a university campus. The comparison proves the accuracy of our analytical models. Moreover, results unveil the key relevance for FC performance of group dynamics in user movements.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

We can identify two main directions in wireless networking research in these years. The first one focuses on evolutionary research aiming at the development of cellular systems of a new generation, with industry-led research projects, mostly sponsored by the European Commission through the 5G Infrastructure Public Private Partnership (5G PPP) [1]. The second focuses on (mostly academic) attempts to identify new architectures for wireless networks, capable of adapting to the environment that must be served by the network. Adaptability can be achieved today through adequate algorithms that exploit information about context, location and user needs, and possibly in the future through artificial intelligence approaches. The latter will leverage on the ability to learn the short-term characteristics of the environment in which the network operates, including user requirements, and continuously optimize the network algorithms accordingly. This second area of research comprises many threads looking at the provision of pervasive, ubiquitous communication services that can be better tailored to the specific needs of groups of end users in a given area at a given time.

Locality of information relevance is naturally paired with localized information exchange. It allows to exploit direct opportunistic communications among end-user devices with a device-to device (D2D) approach [2], without the (or with a minimal) intervention of the network operator resources (but possibly under the network operator control, when under network coverage).

This exploitation of local opportunistic D2D communications to disseminate locally the data that are relevant only in a given area has the advantage of offloading information transfers with local relevance from the bandwidth-hungry infrastructured wireless access and network core transit. It also has the potential for reducing latency in information dissemination. This explains the growing interest that we observe today in D2D approaches and in opportunistic communications.

An interesting example of opportunistic communication approach for the local dissemination of information to end users through D2D connectivity is Floating Content (FC) [3]. FC was conceived to support infrastructure-less distributed content sharing over a given geographic area called *Anchor Zone* (AZ). The objective of FC is to ensure the availability of some content items within the AZ for a







^{*} Corresponding author at: Politecnico di Torino, Italy.

E-mail addresses: gianluca.rizzo@hevs.ch (G.A. Rizzo), vincenzo.mancuso@imdea.org (V. Mancuso), shahzad@ieee.org (S. Ali), marco.ajmone@imdea.org, ajmone@polito.it (M. Ajmone Marsan).

Α	Area of the anchor zone	> 0
\mathcal{A}	Steady state average content availability	[0,1]
$\mathcal{A}(t)$	Average content availability at time t after content generation	[0,1]
AB	Attractor-based	
AP	Attractor point	
AZ	Anchor zone	
$C_x(\cdot)$	Complementary CDF a Poisson process with intensity x	
$f_x(\cdot)$	Probability density function of an exponential r.v. with rate x	
Н	Fraction of active APs containing one node only	[0,1]
Κ	Average number of nodes per active AP	[1, N]
m(t)	Average number of APs with content at time t after content generation	
т	Average number of APs with content in steady state	
Μ	Average number of active APs in steady state	
n(t)	Average number of nodes with content at time t after content generation	
n	Average number of nodes with content in steady state	
Ν	Average number of nodes in AZ in steady state	
р	Probability to jump to a point within AZ	[0,1]
p_{static}	Probability to get content while stopped, because of other nodes jumping	
p_{jump}	Probability to jump in range of a node with content	
p_{stop}	Probability that another node jumps within transmission range	
$P_m(\tau)$	Probability to get content after a jump, s.t. the node leaves AZ at time $ au$	[0,1]
$P_s(\tau)$	Probability to get the content while stopping, s.t. the node leaves AZ at time τ	[0,1]
PJ	Poisson jump	
r	Transmission range of a node	≥ 0
R	Radius of a circular AZ	≥ 0
S	Success ratio (probability to get the content before leaving AZ)	[0,1]
$S(\tau)$	Success ratio, s.t. leaving AZ after τ time units	[0,1]
t	Time since content generation	≥ 0
Т	Average sojourn time of a node in AZ	≥ 0
ν	Velocity of users (∞)	∞
α	Probability to jump within the range of another node	[0,1]
β	Intensity of the Poisson process used for APs	≥ 0
γ	Intensity of the Poisson process used for node arrivals	≥ 0
η	Encounter index	≥ 0
μ	Rate of the stopping time random process	≥ 0
ν	Frequency of encounters between two nodes in AZ	≥ 0
τ	Time since the entrance of a node in AZ	≥ 0

 Table 1

 Notation used in this article.

specified time period by replicating them opportunistically to users which come in contact within the AZ, so that the content items can "float" within the AZ.

Previous works on the performance of the FC service focused on determining the conditions under which a content item floats with high probability [3] and on application-level performance modeling [4]. These works are based on simplifying assumptions for user mobility, and on a simplified model for data exchange. Such models give indications about the potential and limitations of FC, but they offer only first-order predictions of performance in realistic settings. Indeed, important aspects, such as the combined impact of the actual mobility patterns of users, of the communication protocols, of the specific propagation characteristics in the chosen area, of localization accuracy, are difficult to investigate analytically and to evaluate by simulation. Moreover, it is important to understand the impact of the limitations and of the specific features of available protocols for opportunistic terminal-to-terminal communications (such as Bluetooth and WiFi Direct) on the performance of the FC service.

A first experimental study of the operation of FC in a university campus context was presented in [5]. There we described a smartphone app which supports FC and we presented results collected over one week of operation of FC. In the same paper we also presented a simple first attempt to analytical modeling of FC in a campus scenario, which could only be used for the computation of few performance metrics.

In this paper we present a refined version of the analytical model of FC that was introduced in [5], and we show that this new model is capable of capturing the key FC characteristics that were observed in the campus experiment.

The rest of this paper is organized as follows. In Section 2 we provide the system model, as well as some basic details on the operation of FC services. In Sections 3 and 4 we describe the simple analytical models that we developed to predict the FC performance in the class of environments we consider, with and without user clustering. Section 5 presents and discusses analytical results. In Section 6 we describe our experimental setup. Section 7 illustrates some of the results collected during our experiments, and provides a validation of the analytical models. Section 8 discusses some related work, and Section 9 concludes the paper.

2. System model

The notation used in the article will be progressively presented in this section and in the sections dedicated to the system analysis, and it is summarized in Table 1.

We consider a set of wireless nodes carried by users moving (walking) on \mathbb{R}^2 . Users can communicate with each other within a 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*. Every node knows its exact position in space at any time.

We assume that the mobility pattern of each node alternates between time intervals spent moving, and time intervals spent still at a waypoint. This assumption is a good fit for typical features of pedestrian mobility (as well as vehicular). In an outdoor setting, such waypoints could model stops at points of attractions such as bars and restaurants, or pauses at crossroads. In indoor settings, pauses at waypoints model time spent sitting, e.g., at a desk, or time spent within a same room, such as a meeting room, a coffee room, a classroom, and so on. In order to take into account such features of human mobility, we introduce two mobility models, which abstract from the details of the geometry of the space. The first one is the *Poisson Jumps mobility model*. Its formal definition is as follows:

Definition 1 (Poisson Jumps (PJ) mobility model). Consider a region of \mathbb{R}^2 . We assume that nodes arrive in such region of the plane according to a Poisson process with intensity γ . While inside the region, nodes move according to the Random WayPoint (RWP) mobility model with pauses, with velocity $\nu = \infty$, equal for all nodes.

When a node arrives in the region, 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 (its *stopping time*), which we model as a random variable with negative exponential probability density function (pdf) $f_{\mu}(t)$ and rate μ . 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.

In such a mobility model, a sojourn in the given region of the plane is modeled as a sequence of (a random number of) jumps into a set of locations within the region, and a sequence of stops at those locations. Since the probability mass function of the number of stops is geometric with parameter p, and the pdf of the duration of each stop is exponential with parameter μ , the sojourn time pdf is also exponential, with average

$$T = \frac{1}{\mu(1-p)} \tag{1}$$

Note that this model naturally applies to settings where users move "too fast" for content exchange to take place on-the-fly, and where most of the time is spent on a seat (e.g., for our on-campus experiments, in office or in class). Moreover, note that the resulting spatial distribution of nodes is uniform at any time.

Apart from the infinite speed assumption, the main difference between the typical indoor mobility of users and the Poisson Jumps model is that the latter is isotropic (i.e., in any point in space all directions can be chosen with equal probability) and unconstrained (i.e., a user can occupy any location within the area). However, when a "sufficiently large" portion of space is considered inside the region (say, several blocks in a city), even if the possible instantaneous directions of movement are in practice 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, at a macroscopic scale, and for the purpose of modeling content replication and diffusion dynamics within the region, these differences only impact the mean moving time between two waypoints.

One of the main drawbacks of the Poisson Jumps mobility model is represented by the role that the spatial dimension plays in determining the contacts between nodes. Indeed, a well known feature of human mobility patterns, both indoor and outdoor, is that they often take the form of a walk between a sequence of locations, which are therefore attractors for moving nodes [6]. Examples are bars and restaurants or shops for outdoor scenarios, and coffee machines, meeting rooms, classrooms for indoor scenarios. When a node is at the location of an attractor, the spatial component plays little role, because either all nodes are within range of each other, or node connectedness is sufficiently high for a content to be rapidly spread through multi-hop relaying (with no store-carry-and-forward) among all nodes in the vicinity of the attractor location. In order to model the FC performance in such scenarios, we consider a variant of the Poisson Jump mobility model, the *Attractor-Based mobility model*.

Definition 2 (Attractor-Based (AB) mobility model). Consider a region of \mathbb{R}^2 , with area *A*. We assume that nodes arrive in such region of the plane according to a Poisson process with intensity γ . Let us assume that in the region we have *attractor points* (APs), spatially distributed according to a two-dimensional Poisson point process of intensity β . Nodes inside the region move according to the Poisson Jumps mobility model in which the only locations allowed for waypoints are either the attractor points or the points at the border of the AZ. When a node is at an attractor point, and jumps inside the region, the candidate points for the next waypoint are the remaining attractors (i.e., at the end of a stopping time a node cannot remain at the same waypoint).

The average sojourn time in the AZ remains as in the PJ model, given by (1). Note that for $\beta \rightarrow \infty$ the AB mobility model coincides with the PJ mobility model, as stated in the following proposition.

Proposition 1. If the intensity β of the AB model diverges, the AB model converges in probability to the PJ model.

Proof. Note that the difference between the PJ and AB mobility models is that the PJ model allows jumps to any point within the AZ, while the AB model imposes jumps to attractor points only. In the AB model, the probability that a disk or radius $\rho > 0$ within the AZ contains at least one attractor is $1 - e^{-\beta \pi \rho^2}$. Such probability converges to 1 if $\beta \to \infty$ for all positive values of ρ . Therefore, if β diverges, any disk contained in the AZ contains at least one attractor. If now we let $\rho \to 0^+$, we obtain that any point in the AZ becomes an attractor with probability 1. Therefore, a node that jumps can select any point in the AZ as a possible destination, which is what assumed by the PJ model. \Box

2.1. The Floating Content service

Floating Content (FC) is an information sharing paradigm that allows the implementation of a class of infrastructure-less services based on opportunistic wireless terminal-to-terminal communications. We assume that at time t = 0 a "seeder" node produces a content item which is of interest for users within a limited geographical area, the AZ, for a given interval of time. Therefore, the content item is generated inside the AZ. When nodes move outside of the AZ, they delete their own copy of the content item [3]. Whenever a node having a content item in the AZ comes in contact with some other node that does not have it, the content item is replicated through opportunistic message exchanges, as shown in Fig. 1. In this way, content items *float* within the AZ, i.e., they are available on a set of nodes that move within the AZ. The set of nodes carrying each content item varies over time, even after the node that generated the content item has left the AZ.



Fig. 1. Operation of a Floating Content service.

Through this geographically constrained opportunistic replication mechanism, a given content item is stored probabilistically in a spatial region, typically without the support of fixed infrastructure, and it is made available to users traversing the AZ through opportunistic exchanges with nodes in the AZ.

Previous works showed that content items float within the AZ over time with high probability provided a *criticality* condition is met, which accounts for the average number of nodes in the AZ (denoted by *N*), the average contact rate between any two nodes (ν), and the average node sojourn time in the AZ, expressed as in (1) [3,7]. In our case, the criticality condition is as follows:

$$\eta = N\nu T = \frac{N\nu}{\mu(1-p)} > 1$$
(2)

We denote η as the *encounter index*, and we will show how to compute it under different node mobility models.

The criticality condition on the encounter index η is a sufficient condition for content to persist indefinitely over time within the AZ. Indeed, when the criticality condition is satisfied, the expected lifetime of a content item is infinite under the fluid limit approximation of [3]. In practice, this implies large average lifetimes for content item.

However, the sole condition of floating does not provide service guarantees. In fact, a content item needs to be *available* for new nodes to obtain when they enter the AZ, and it needs to reach the nodes transiting through the AZ with suitably high probability [4]. More in detail, the following definitions are key to the formulation of performance metrics for FC services.

Definition 3 (Availability). The availability of a content item measured *t* time units after its generation is denoted by A(t) and is defined as the ratio between the average number of users in the AZ with content at time *t* and the average number of users in the AZ at time *t*.

Definition 4 (Steady state availability). The steady state availability of a content item is denoted by A and is defined as the ratio between the average number of users inside the AZ holding a copy of that content item and the average AZ population in steady state.

Definition 5 (Success probability in time *t*). The success probability of a content item in a sojourn time of duration *t* time units is denoted by S(t) and is defined as the probability that a node obtains a content item in a time not greater than *t*.

Definition 6 (Success probability). The success probability of a content item is denoted by S and is defined as the probability that a node obtains a content item before leaving the AZ.

3. FC performance under the Poisson Jumps model

In this section, we present a novel analytical model for the computation of the main performance parameters characterizing FC services in areas like a university campus or a large office building. This model is an extension of the one presented in [5]. The analysis is based on the PJ mobility model. For all proofs in this section, please refer to Appendix A.

As observed in experimental data, we assume the mean stop time to be much longer than the average time for content item transfer. Thus, transfers can be considered instantaneous for all practical purposes. We also consider stop times to be long enough to allow retransmissions in case of failure. Hence, we can neglect the effect of communication errors.

The AZ is assumed to be circular, with radius *R*. We assume moreover that $R \gg r$; otherwise, direct communications between nodes in the range of the seeder would be enough to spread the content, hence making it useless to have an AZ.

As we mentioned, one of the main performance parameters for FC is the probability for a user to get a content item during its sojourn in the AZ, since this quantity measures the efficiency with which the FC service makes information available to intended users. In order to derive such performance parameter, we introduce the following results.

Lemma 1 (Two-nodes contact rate in the PJ model). The frequency with which two nodes come in contact within the AZ in the PJ mobility model is given by $v = 2\mu\alpha$, where α is the probability for a jumping node to land in range of a given node, which is computed as $\alpha = \frac{r^2}{R^2}$ under the PJ model assumptions.

In what follows, we consider an AZ in steady state. That is, we assume that any transient in the temporal evolution of node population due to initial conditions (e.g., system starting from a given number of users, at some point in time), as well as any transient in the temporal evolution of the population of nodes with content in the AZ due to initial content seeding, are extinguished. In such steady state conditions, we use N to denote the average number of nodes in the AZ, and n for the average number of nodes with content item.

Using Little's theorem, the average population in the AZ is $N = \frac{\gamma}{\mu(1-p)}$, whereas the following theorem derives an expression for the mean number of nodes with content in the AZ.

Lemma 2 (Steady state availability in the PJ model). Consider an AZ in steady state where nodes move according to the PJ model. The criticality condition (2) becomes

$$\eta = \frac{\nu N^2}{\gamma} > 1 \tag{3}$$

If (3) *is satisfied, then the following relations hold:*

$$\mathcal{A} = 1 - \frac{1}{\eta} \tag{4}$$

$$n = N\left(1 - \frac{1}{\eta}\right) \tag{5}$$

The more the encounter index—that, as shown by (3), is proportional to the ratio between the overall contact rate in the AZ and the arrival rate—is bigger than 1, the less likely is the content to disappear during a given time interval due to the natural random fluctuations in the population of nodes with content, which occur in a system based on opportunistic content replications.

Lemma 2 allows us to derive the mean amount of nodes with content within the AZ in stationary conditions when the critical condition is satisfied. Availability is tightly related to success probability, the main performance metric for FC. Indeed, in order to achieve a given percentage of users exiting the AZ with content, one of the requirements is to have a "sufficiently large" percentage of nodes in the AZ with content, in order to bring the likelihood for the exiting node of having met at least one node with content during its sojourn in the AZ to the desired level.

The following theorem derives an expression for the success probability under the PJ model, and provides the relation between success probability and availability.

Theorem 1 (Success probability in the PJ model). Consider a content item floating within a circular AZ with radius R, in which nodes move according to the PJ model. In the stationary regime, when the criticality condition (3) is satisfied, the probability for a node entering the AZ to get the content item during its sojourn in the AZ is given by

$$S = \frac{p_{stop}}{1 - p\left(1 - p_{stop}\right)} \tag{6}$$

where p_{stop} is the probability of getting the content item during the stop time that follows a jump:

$$p_{stop} = p_{jump} + (1 - p_{jump}) p_{static}$$

where p_{jump} is the probability of getting the content as soon as a node jumps in a new location, while p_{static} is the probability of getting the content during the stop time, thanks to other nodes jumping in locations within range of the given node.

These quantities are given by

$$p_{jump} = 1 - e^{-n\alpha} \tag{7}$$

$$p_{static} = \int_{0}^{+\infty} f_{\mu}(\tau) e^{-\alpha \tau \frac{\gamma}{1-p}} \sum_{j=1}^{+\infty} \left(1 - \eta^{-j}\right) \left(\frac{\alpha}{1-\alpha}\right)^{j} C_{(1-\alpha)\tau \frac{\gamma}{1-p}}(j) \, d\tau \tag{8}$$

where $f_{\mu}(\cdot)$ is the pdf of the negative exponential random variable with rate μ describing the stopping time, and $C_{(1-\alpha)\tau} \frac{\gamma}{1-p}(\cdot)$ is the complementary CDF of a Poisson process with intensity $(1-\alpha)\tau \frac{\gamma}{1-p}$. We recall that $\alpha = \frac{r^2}{R^2}$ under the assumptions of the PJ model.¹

In practical settings, transients (initial and final) as well as fluctuations in the population of users impact the performance of the FC service in a measure which depends on content item lifetime and on many other system parameters. Nonetheless, Theorem 1 provides a reasonable estimate of the success probability of a given content item in practical settings, and it allows estimating the impact of various system parameters on performance.

A similar result can be derived for the probability of obtaining a content item while spending exactly τ time units within the AZ, as expressed in the following theorem.

Theorem 2 (Probability of getting the content item during a sojourn of τ time units in the AZ). With the same assumptions as Theorem 1, when stopping times are exponentially distributed and the criticality condition (3) is satisfied, the probability for a node in the AZ to possess the content τ time units after its ingress in the AZ, is

$$\mathcal{S}(\tau) = P_{s}(\tau) + (1 - P_{s}(\tau))P_{m}(\tau) \tag{9}$$

where $P_{s}(\tau)$ is the probability of receiving the content while not moving, conditioned to a sojourn time of duration τ :

$$P_{s}(\tau) = e^{-\alpha \tau \frac{\gamma}{1-p}} \sum_{j=1}^{+\infty} \left(1 - \eta^{-j}\right) \left(\frac{\alpha}{1-\alpha}\right)^{j} C_{(1-\alpha)\tau \frac{\gamma}{1-p}}(j)$$
(10)

¹ Note that $(1 - \alpha)\tau \frac{\gamma}{1-p}$ is the steady state number of jumps of nodes that land within AZ and off the reach of other nodes in an interval τ .

while $P_m(\tau)$ is the probability of receiving the content as a result of jumping to a new location:

$$P_m(\tau) = 1 - e^{-\mu\tau p_{jump}}$$
where p_{jump} is given by (7). (11)

4. FC Performance under user clustering

In this section we modify the previous derivation in order to obtain the FC performance parameters when the mobility model accounts for the presence of attractors. For proofs in this section, we refer readers to Appendix B.

Lemma 3 (Two-nodes contact rate in the AB model). Under the AB model, the frequency with which two nodes come in contact within the AZ is $v = 2\mu\alpha$, with

$$\alpha = \frac{1}{\beta \pi R^2} + \left(1 - \frac{1}{\beta \pi R^2}\right) \frac{r^2}{R^2}$$

In what follows we denote those APs hosting at least one node as *active APs*. Let *M* and *m* denote, respectively, the average number of active APs and the average number of active APs in which nodes have the content item in steady state, respectively.

Lemma 4 (Average availability in the AB model). Consider an AZ in steady state with an average number of nodes equal to N. Under the AB model, the criticality condition becomes

$$\eta = \frac{K}{H} \frac{\nu N^2}{\gamma} > 1 \tag{12}$$

and when (12) is satisfied the steady state availability is given by:

$$\mathcal{A} = \frac{n}{N} = \frac{m}{M} = 1 - \frac{1}{\eta} \tag{13}$$

where H is the fraction of active APs hosting just one node:

$$H = \left(1 - \frac{1}{\beta \pi R^2}\right)^{N-1} = \frac{\left(1 - \frac{M}{\beta \pi R^2}\right)}{\left(1 - \frac{1}{\beta \pi R^2}\right)}$$
(14)

and $K = \frac{N}{M}$. If the criticality condition holds we also have

$$M = \beta \pi R^2 \left[1 - \left(1 - \frac{1}{\beta \pi R^2} \right)^N \right]$$

$$m = M \left(1 - \frac{1}{\eta} \right)$$
(15)
(16)

Note that, as $0 \le H \le 1$, and since $K \ge 1$, when nodes cluster according to the AB model, the criticality condition gets looser than the one for the PJ model, when all other system parameters are the same. Such intuitive result implies also that the mean availability under the AB model is higher than in the setup without clusters.

The expressions for the AB model are compatible with the ones obtained for the PJ model. Indeed, for $\beta \to \infty$ we have that $H \to 1$ and $M \to N$, so that $\frac{K}{H} \to 1$. Under the same conditions, $\alpha \to \frac{r^2}{R^2}$ and therefore all the equations derived here for the AB model become the ones derived for the PJ model (with n = m) when the density of attractors becomes sufficiently high.

Theorem 3 (Success probability in the AB model). Let us consider a circular AZ in which nodes move according to the AB model. In the stationary regime, when the criticality condition (12) is satisfied, the probability for a node entering the AZ to get the content item during its sojourn in the AZ is given again by (6), where p_{jump} is given by

$$p_{jump} = \frac{m}{\beta \pi R^2} + \left(1 - \frac{m}{\beta \pi R^2}\right) \left(1 - e^{-m\frac{r^2}{R^2}}\right)$$
(17)

 p_{static} is again given by (8), with the expression for α given by Lemma 3.

Note again that, for $\beta \to \infty$, Theorem 3 satisfies the regression test to the result based on the PJ model (Theorem 1). This implies that by varying β while keeping everything else constant, Theorem 3 allows modeling the FC performance in scenarios with various degrees of user clustering and of inter-cluster content diffusion.

Note that the case with no inter-cluster content diffusion (modeling, for instance, setups where clusters are far away, or are indoor rooms separated by walls opaque to transmissions), can be easily derived from Theorem 3 by letting $r \rightarrow 0$.

Finally, under the AP model, when stopping times are exponentially distributed, it is easy to verify that the probability for a node in the AZ to possess the content *t* seconds after its ingress in the AZ is still given by Theorem 2 with the expressions for α and p_{jump} of Theorem 3.

5. Model results

In this section, we exploit the analytical models that were presented in the previous section to characterize the performance of FC under either the PJ or the AB mobility models. Later, we will validate the analytical results with experimental results.



Fig. 2. Success probability and availability for one content item vs the ratio $\alpha = r^2/R^2$, for three different values of *p*. Vertical lines indicate the minimum values of α for which the criticality condition is satisfied. PJ model. $\gamma = 0.1$ nodes/s, $1/\mu = 600$ s.



Fig. 3. Success probability and availability for one content item vs encounter index η , for three different values of p. PJ model. $\gamma = 0.1$ nodes/s, $1/\mu = 600$ s.

5.1. Numerical results

5.1.1. PJ model

We start by considering the PJ mobility model and we assume that a new mobile terminal arrives to the AZ on average every 10 s ($\gamma = 0.1$ arrivals per second), and that users stop at wayponts for an average time of 10 min ($1/\mu = 600$ s).

In Fig. 2 we plot the success probability and the availability for a content item versus the ratio $\alpha = r^2/R^2$, for three different choices of the probability *p* of jumping within the AZ, which, according to (1), imply three different values for the mean sojourn time in the AZ, *T*. The vertical lines indicate the minimum values of α for which the criticality condition is satisfied for the three different values of *p*.

When the average sojourn time in the AZ is longer, the average number of terminals in the AZ is larger (due to Little's result), so that a smaller transmission radius is sufficient to achieve the amount of content exchanges per unit time, which are able to sustain high values of success probability and availability. Availability values are initially higher than values of the success probability, but an inversion is observed for growing values of α . This inversion happens earlier for lower values of the sojourn time in the AZ. The figure also shows that availability and success probability in steady state are not far apart.

In order to abstract the analysis of the FC behavior from the geometrical characteristics of the system (in particular, r and R), we plot success probability and availability versus the encounter index η , like in Fig. 3. We can see that availability only depends on η , as stated in (4), and that also the success probability is mostly function of the encounter index, with relatively minor variations for different values of p. We again see that the availability is larger than the success probability for smaller values of the encounter index, while the converse is true for higher values of the encounter index. Intersections happen earlier for lower values of p (hence of T). Counter-intuitively, with increasing mean sojourn time in the AZ, the success probability tends to decrease.

In order to better understand the behavior of the curves in Fig. 3, we have plotted in Figs. 4 and 5 the parameters p_{static} (i.e., the probability for a node to get a content item by having other nodes jumping into the node range during a stopping time) and p_{jump} (that is, the probability for a node to get a content item by jumping into the range of at least one other node with that content item). From the plots we see that, for the range of values of system parameters considered in Fig. 3, the contribution to success probability due to p_{static} is practically negligible with respect to the one due to p_{jump} (note the multiplier on the *y*-axis in Fig. 4). Moreover, quite counter-intuitively, we see that both probabilities sensibly decrease with increasing *p* (and hence with the node mean sojourn time in the AZ).

Another interesting aspect that we can observe is that, despite this marked dependency on *p* of both probabilities, the overall success probability (Fig. 3) does not exhibit a drastic change with varying *p*, decreasing only slightly with increasing *p*. Indeed, by keeping constant the encounter index and increasing *p* we increase the mean sojourn time in the AZ, hence the mean number of nodes in the AZ, hence the mean overall contact rate in the AZ. To keep the same value of the encounter index, we have to modify the system geometry, reducing the ratio r^2/R^2 . This has no effect on availability, as we have seen. However, in order to attain the same value of availability with the same value of encounter index (hence of contact rate) while increasing *p* (hence spending on average more time in the AZ), both p_{static} and p_{jump} have to decrease. This is precisely what can be observed in Figs. 4 and 5. The same happens for the overall success probability, since the growth in the average time spent in the AZ roughly balances the decrease in these probabilities, as it does for availability.



Fig. 4. p_{static} vs encounter index η , for three different choices of the probability p. PJ model. $\gamma = 0.1$ nodes/s, $1/\mu = 600$ s.



Fig. 5. p_{jump} vs encounter index η , for three different choices of the probability p. PJ model. $\gamma = 0.1$ nodes/s, $1/\mu = 600$ s.



Fig. 6. Success probability and availability vs encounter index η , for three different values of mean stopping time (hence of the mean sojourn time in the AZ). PJ model.

If we fix the value of p, and we change the average node sojourn time in the AZ by increasing the value of the stopping time, we obtain the results in Figs. 6–8. From those we see that success probability, availability, and p_{jump} are invariant with respect to changes in the average stopping time value. On the contrary, p_{static} significantly decreases for increasing average stopping time.

5.1.2. AB model

In the case of the attractor-based model, the expression for the encounter index becomes as in (12), and the minimum value of η as a function of r is clearly reached when r = 0, and it is given by

$$\eta_{\min} = \frac{K}{H} \frac{2\mu N^2}{\gamma \beta \pi R^2}$$

Fig. 9 shows how the encounter index varies with α , hence with *r*, in the case of three different values of the intensity of the Poisson process generating APs in the AZ, hence of mean number of nodes per AP (5, 0.5, 0.05). When enough nodes are present at each AP, with decreasing transmission range, the encounter index stabilizes at a value well above one. In this case, even without inter-cluster diffusion, the content fluctuates indefinitely over time. When we increase the number of APs while keeping fixed the mean number of nodes, so that we decrease the degree of node clustering, we see that the encounter index decreases. When the mean number of nodes per AP is low enough to have very low or no clustering, diffusion effects (i.e., jumping in range of another node, rather than jumping exactly at its location) start playing a key role in the overall FC performance. Thus, for low enough values of α and hence of transmission range, the encounter index is below 1 and content does not float.

The corresponding values for success probability and availability are shown in Fig. 10. As expected, the more the encounter index is far from one, the closer are the main performance parameters to their maximum values.



Fig. 7. p_{static} vs encounter index η , for three different values of mean stopping time (hence of the mean sojourn time in the AZ. PJ model. T = 1800 s, $\gamma = 0.1$ nodes/s, p = 0.5.



Fig. 8. p_{jump} vs encounter index η , for three different values of mean stopping time (hence of the mean sojourn time in the AZ). PJ model.



Fig. 9. Encounter index vs α , for three different values of number of APs in the AZ, hence of mean number of nodes per AP. N = 150, R = 300 m, T = 1500 s, $\gamma = 0.1$ nodes/s, $\mu^{-1} = 600$ s, p = 0.6. AB model.

In absence of inter-cluster content diffusion (i.e., for $r \rightarrow 0$), for a given total number of nodes in the AZ *N*, arrival rate γ , and jump frequency μ , the maximum of the mean number of APs in the AZ that make the content float indefinitely can be obtained by using the criticality condition (12). Using $\lim_{r\to 0} \alpha = \frac{1}{\beta \pi R^2}$ and $\nu = 2\mu \alpha$, and considering that the average number of APs in the AZ is $\beta \pi R^2$, the criticality condition imposes that

$$\lim_{r \to 0} \eta = \frac{K}{H} \frac{\frac{2\mu}{\beta \pi R^2} N^2}{\gamma} > 1 \Rightarrow \beta \pi R^2 < \frac{K}{H} \frac{2\mu N^2}{\gamma}$$

The above is an implicit equation because both *K* and *H* depend on $\beta \pi R^2$.

In Fig. 11 we plot the ratio between such number of APs in the AZ and the mean number of nodes in the AZ. It is evident that such ratio is practically constant for varying N. Intuitively, the larger the node sojourn time T in the AZ, the higher such ratio, as nodes have more time to come in contact with a node with content by jumping among APs.

6. Experimental setup

In this section we describe our experience with FC services in a large university campus scenario. The experimental FC services were running on *Floaty*, an Android application we developed and tested in the Leganés campus of the University Carlos III of Madrid, Spain.



Fig. 10. Success probability and availability vs *a*, for three different values of number of APs in the AZ, hence of the mean number of nodes per AP. N = 150, R = 300 m, T = 1500 s, $\gamma = 0.1$ nodes/s, $\mu^{-1} = 600$ s, p = 0.6. AB model.



Fig. 11. Maximum ratio between the mean number of APs and of nodes in the AZ, beyond which the criticality condition is not satisfied in the regime with no inter-cluster diffusion. $R = 300 \text{ m}, \mu^{-1} = 600 \text{ s}$. AB model.

6.1. Floaty: a Floating Content mobile app for Android

In order to perform our tests we implemented the FC service on an Android smartphone application named *Floaty*. Floaty implements opportunistic communications over Bluetooth, which we selected because it is presently available in almost all smartphones in the market. WiFi Direct was the possible alternative to Bluetooth, but it was discarded because it typically consumes much more energy [8], and poses severe technical problems and security issues, due to the lack of simple user authentication modes.

The content items generated by Floaty are lightweight. They are composed of an identifier of the device that generated the content item (its MAC address and Bluetooth name), a timestamp (the time at which the content item was generated), and a sequence number. As a consequence of the small size of each content item, during the whole experiment, a single Bluetooth message was sufficient to transfer all of the content items that were transmitted from one node to another. This allowed us to neglect the effect of content item size over transfer time, and, in general, over the performance of the FC service.

When a device running Floaty is in the AZ, the app generates a new content item every 15 min, and content items are transferred to all the other nodes that come within the radio transmission range of such device, and which are running the same app. A content item does not expire until the end of the day when it was generated, so every day we generate a fresh population of content items. Note that the availability of several content items generated by each user during its sojourn inside the AZ allowed the test of FC services performance with a fine granularity over time and space in the experimental setting.

In order to detect the presence of Floating Content items, every 60 s Floaty performs a Bluetooth scan, and builds a list of available Bluetooth peers running the same app.² Whenever the list of peers is nonempty, the app chooses one of them and tries to establish a Bluetooth connection with it. If it succeeds, the two peers exchange all content items they store. However, due to Bluetooth limitations, the app connects to one peer at a time, and spends a few seconds to transfer the available content items between connected peers (up to 12 s in our experiments, most of which due to Bluetooth connection protocols, rather than to data transfer, as we have experimentally verified). Therefore, it is important to optimize the choice of the peers to improve content items diffusion in dynamic settings. To this end, Floaty ranks peers in radio range according to the time of the last content item exchange, starting from the oldest ones, so that the peers that never met before have absolute precedence.

For the FC mechanism to operate correctly, the most relevant information is whether a user is inside the AZ, rather than its exact position in space. Therefore, instead of exploiting GPS signals or the like, the Floaty app just makes use of a coarse localization method based on WiFi signals. Specifically, to decide whether a node is inside the AZ, every 60 s the Floaty app scans for signals of surrounding access points. If it detects at least one reference access point (any of the ones present on campus), then it assumes to be inside the AZ. Therefore, our experiment covers a slightly larger area than the university campus. However, being the campus quite large, and being the access points located inside buildings, the difference is negligible.

² This is possible because Bluetooth allows to announce running services.



Fig. 12. Behavior of users over a day (average over 5 days). Data refers to number of nodes, number of contacts, arrivals and departures counted in intervals of 15 min.

All scanning intervals and content items generation intervals were chosen so as to achieve a reasonable tradeoff between the performance of FC services and the energy consumed by the devices. In particular, we quantified the average energy consumption for a test set of smartphones in \sim 5% battery discharge per hour. Generating content items more frequently or scanning WiFi and Bluetooth on a subminute basis would have made terminals incur excessive battery costs. However, we experimentally evaluated our test set of devices using different scanning and content item generation intervals. As a result, we found that, given the slow dynamics of the campus environment, the performance of FC services was not benefitting much from higher scanning frequencies or more frequent content item generation.

Floaty logs all events related to content items transfer between nodes, including timestamps, connection durations, lists of transferred content items. The app also logs all events relating to AZ ingresses and egresses, as well as all failed message transfers. These logs are periodically uploaded to a remote server.

6.2. The experiment

After having tested Floaty with a reduced test set of terminals and users, we advertised the experiment on campus, recruiting students, researchers, and professors to download and install Floaty on their smartphones and tablets. In particular, we had access to the Computer Science building. In total, among the people frequenting the building, 62 volunteered to download and install Floaty on their personal devices. However, only 48 devices produced valid logs. Moreover, a tiny fraction of these users never exchanged any content item during the whole duration of the experiment. The experiment lasted 5 days (Monday to Friday) from 8 A.M. to 6 P.M. During the whole experiment, 1117 content items were generated.

As we will show in Section 7, we verified that users had a peculiar mobility pattern, i.e., they moved between home and the university, and, once on campus, they had the tendency to form stable groups and spend long intervals in the same place (as expected for students attending classes). The result of such mobility patterns is scarce spatial interaction with most of the other users. Indeed, this is a trivial but important aspect of the performance of FC services. In this regard, the experiment confirmed that infrastructure-less FC is a reasonable choice even for applications meant for users with reduced spatial interactions. However, for our evaluations, we considered only those content items that got a chance to be replicated at least once, which consists in a population of 923 content items.

7. Experimental results and model validation

In this section we report the main results of our experiments. We first use the collected logs to characterize the behavior of users participating in the experiments, in terms of mobility, connectivity, and generation of content items. Afterwards, we show the performance figures achieved by the FC service, mainly in terms of availability and success probability of content items.

These experimental results will on the one hand justify the choices we made in the development of analytical models, and on the other allow the validation of the model predictions.

7.1. Characterization of user behavior

During the experiments, users dynamically enter and exit the AZ and generate content items only when they are within the AZ, so that the number of nodes present in the AZ, as well as the number of content items floating in the AZ, changes over time. Fig. 12 illustrates the dynamics of number of nodes and contacts observed in the AZ over the day, averaged over the logs collected during the five days of experiments. For ease of presentation, arrivals, departures, number of nodes, and number of contacts are counted with a time resolution of 15 min.

From the figure, it emerges that the average number of nodes does not change drastically over the day, although it is possible to identify five regions in the graph: (i) the leftmost part of the curves accounts for the starting of a new day, when arrivals are more numerous than departures (8 A.M.–10 A.M.); (ii) from 10 A.M. to 2 P.M., the number of nodes in the AZ changes very little, after which (iii) the curve shows a drop, corresponding to lunch time (2 P.M.–3 P.M.); (iv) from 3 P.M. to 5 P.M. the number of nodes is again practically stable, while (v) after 5 P.M. there is a prevalence of users leaving the campus and thus the AZ.

The number of contacts between nodes does not follow exactly the same trajectory as the number of nodes, because users within the AZ often move in groups, attend classes, meet in shared areas and cafeterias, so that the connectivity pattern is more irregular over time. In particular, note that nodes present in the AZ at the beginning and at the end of the day do not experience many contacts, which is a



Fig. 13. Connectivity graph for different times of the day. Each point corresponds to a node, and a link represents one or more contacts between two nodes during the considered hour of the day.



Fig. 14. Empirical CDFs of user sojourn time and of time to get a content item, and probability of getting the content vs time spent inside the AZ, for the PJ model (Theorem 2).

symptom of scarce mobility of users arriving early and/or leaving late (probably these are professors or researchers sitting in their offices, where they are under campus WiFi coverage). In general, Fig. 12 shows that contacts are, on average, not very frequent. For instance, about 16 users experience in total 33 contacts at 1 P.M. (i.e., between 1 P.M. and 1:15 P.M.), which is a symptom of limited mobility on the time scale of minutes.

Overall, the results in Fig. 12 show that the assumption of steady-state, that we used in the development of analytical models, is roughly met for a significant portion of our experiments.

Fig. 13 gives more insight on the occurrence of contacts between users. The figure depicts two snapshots of the connectivity graph between users in the AZ. Every link between two users means that those users had a contact in the considered time interval, in the first day of the experiment. Clearly, connectivity exhibits drastic changes over the day, which can affect the performance of FC services in terms of users that can be reached, and probability of the content items floating within the AZ. Notably, the figure unveils that one or two components are practically sufficient to cover the full set of users experiencing contacts, so that we can infer a *clusterization* of user contacts in which two clusters are practically enough to describe the process. This observation led to the development of the AB mobility model. However, the experiments do not indicate what is a reasonable number of "attractors". The observed behavior could indicate either that there are only 2 or 3 attractors that can be visited, or that people move in groups, and only 2 or 3 attractors have users at any point in time, while others remain empty.

Our choice of modeling nonuniform user distributions in space through APs is one of the innovative features of our modeling approach.

Our logs also show that only 85% of the contacts generate a successful transfer of content items. This behavior is due to users trying to exchange content items while actually moving, e.g., walking through a corridor, and, in general, when there is not enough time to complete the Bluetooth connection procedure and the actual message exchange. Indeed, considering that the rate of unsuccessful content transfers is low, and that the number of contacts logged in our experiments is limited and shows clusterization effects, we infer that in our experiments users tend to move quickly from a place to another and then spend long intervals at their destination (as expected for students attending classes on campus). From our experimental results we can estimate the average stopping time duration to be slightly less than one hour (58 min and 14 s).

Again, the observation of these behaviors motivated the development of our models, which only account for information exchange during the time users spend still at a waypoint.

To further characterize the mobility pattern of users, we plot the CDF of the sojourn times of the users within the AZ in Fig. 14. From the figure, it emerges that about 50% of the users spend 2 h or less inside the AZ before leaving (although they possibly come back later, e.g., after lunch). The average sojourn time in the AZ can be estimated around 1 h and 40 min. The figure also reports the CDF for the time to get a content item once a node enters the AZ, or after a new content item is generated while a node was already in the AZ. Notably, most of the content items are replicated within a few tens of minutes (70% of those nodes that get a given content item do it within 40 min), i.e., much faster than the typical time spent by a node inside the AZ. Therefore, we can conclude that, notwithstanding the



Fig. 15. Empirical CDFs of content item lifetime, conditioned to content item being replicated at least once, and dead before 17.00.



Fig. 16. Mean content item availability over content item lifetime; experiments and model, for different number of seeders.

limited frequency of contacts, the mobility of users is still suitable to make the content items float within the AZ. We empirically validate this statement by analyzing the floating behavior of content items in the next subsection.

7.2. FC performance evaluation

In Section 7.1 we have already mentioned the fact that the time to get a content item is typically short with respect to the time spent by a user in the AZ, and that users exhibit scarce mobility on short timescales. However, from the perspective of an application using FC, floating lifetime of a content item is one of the most studied features of the service, perceived as one of the main measures of its feasibility. Indeed, in a real scenario, all content items naturally disappear at some point in time, not only because of day/night mobility patterns, but also because of stochastic fluctuations in population density and in mobility pattern. Among all content items generated over the five days of the experiment and replicated at least once, only 5% died out before the end of the day. For those content items that do get replicated, but do not reach the end of the day, Fig. 15 shows that the vast majority disappears in the early stages of the diffusion process, while the number of Floating Content item replicas has not reached a critical mass. Conversely, when the number of replicas gets *sufficiently high*, the replication mechanism proves efficient enough to compensate content item replicas lost due to users leaving the AZ.

This observation allows us to compare our experimental data with results from available models. For mobility patterns such as the ones in our experiment, with relatively few "on-the-fly" content item exchanges, from [9] the mean number of nodes in range of a given node should be larger than 1.19 for a content item to float indefinitely with high probability. But despite the very low content item mortality in our experiments, in our setting the mean number of nodes in range has been 0.612, well below the criticality condition.

The net result of the content item dynamics described above is that the average availability of content items grows over time, counting from generation time, as shown in Fig. 16 for the first 50 min of content lifetime.

Observing the experimental results for the mean content availability depicted in Fig. 16, we notice that while the content items float, only about half of the users in the AZ get a copy of the Floating Content items within 50 min. To validate the hypothesis that also this behavior is due to the occurrence of clusters, we parse once again our logs to build the following functions of time: the availability of content items replicated at least once, the one of content items replicated at least twice, and the one of content items replicated at least three times. Furthermore, we build the availability of pairs and triples of content items, i.e., the availabilities obtained by considering the joint availability of respectively two or three content items generated roughly at the same time. In this way, we simulate the presence of multiple *seeders* for a same content item. The resulting curves confirm the non-uniformity in user distribution in space. In fact, multi-seeder curves exhibit high availability levels, while little difference can be noticed between the singe-seeder curves. In practice, since seeders are chosen at random among all users present in the AZ at the time of content item generation, having more seeders improves the probability to reach more clusters within a given time from content item generation. In contrast, if the initial replicas reside in users



Fig. 17. Experimental CDF of success probability, for different number of seeders.

in range of each other, the effect on availability is negligible. The little difference between the case with two seeders and the one with three seeders tells us that the mobility pattern in our experimental settings gives rise to a clustered structure that could be modeled with as few as two clusters.

This also implies that the system behavior can be described with the AP mobility model using just few attraction points.

Indeed, the model results presented in Fig. 16 show that accurately modeling the observed behavior is not easy. Looking at the curves with 3 seeders, we see that the AB model matches well the short term availability (the first 20 min after generation), while the PJ model provides accurate estimates in the second 30 min of the experiment. This may be due to the fact that in a university environment there is little mixing among groups of students, since the same group moves from one class to another. Using the prediction of the AB model in the first 25 min and those of the PJ model for the second part of the interval provides quite an accurate overall estimate of the content item availability.

To complete the performance evaluation of FC services in our experiments, we finally report results on the success probability. This is a measure of the effectiveness in accessing the content items stored (floating) in the AZ. As reported in Fig. 17, the success probability can be estimated via the success ratio, i.e., the ratio between the number of users that obtained a content item and the total number of users in the AZ during the content item lifetime. Users who entered the AZ more than once during the content item lifetime are counted multiple times. The figure shows the cases of one, two or three content seeders. Specifically, the figure reports the empirical CDF of the success probability computed over all the 923 content items considered for the experiment statistics. For the case of one single seeder, the peculiar mobility pattern of the users does not allow high success probability with high probability. In fact, half of the content items have a success probability of 0.5 or less, and just 10% of the content items exhibit a success probability higher than 80%. This result was somehow expected, since we knew from Fig. 16 that the average availability over time does not reach values much higher than 50%. With more seeders, both availability and success probability are boosted. In particular, with two seeders, 65% of the content items are replicated to all users present in the AZ during the lifetime of content items. With three seeders, all users in the AZ, independently from the time spent in the AZ, receive a copy of more than 90% of the content items.

In this paper, we have derived performance metrics for FC. Whether such performance is suitable or not to allow for the diffusion of FC services is out of the scope of this paper, and it mostly depends on the requirements of the specific application to be implemented on top of FC. However, with our analytical and experimental results we have unveiled that, if required by the application, using just a few independent seeders can dramatically increase both availability and success probability, so as to reach the needed performance level. For instance, in our campus experiments, no more that three seeders are needed to make FC extremely effective.

8. Related work

Concepts similar to Floating Content have appeared in the literature with different names. In [10], Castro et al. present a very similar concept called hovering information. The focus of hovering information is sharing content with spatio-temporal locality constraints, using pure ad hoc communications, like FC. The FC concept was first coined in [3], where Hyytia et al. came up with the criticality condition we have used in our work, under which a content can float indefinitely over time. In [11], Ott et al. validated the analytical results presented in [3] and showed the feasibility of Floating Content under the presence of a modest number of nodes in a network by doing extensive simulations. In [4] we considered the performance of context-aware applications that use FC as a communication service. A simple analytical model is presented for computing success probability for two different kinds of context-aware applications under a simple Random Direction mobility model. By doing extensive simulations, we proved that despite its simplicity, the analytical model is capable of accurately predicting the performance of FC for the considered applications. In [12] we presented the performance of contextaware applications using FC under different mobility models. With extensive simulations, we showed that by tuning the anchor zone radius, high success probability can be achieved under a variety of mobility models. In [13], Hagihara et al. proposed a delivery control method for FC called PFCS (Proportional control for Floating Content Sharing). The authors assume that there are multiple different content items present in the network and replicating all of the content items using FC may lead to overloaded wireless network and/or storage capacity of nodes. In order to avoid such situations, PFCS controls the number of copies of a particular content item (called possession ratio) within an anchor zone. By doing a stability analysis, the authors derived a condition called as stability condition in which the possession ratio can be controlled to a target value.

The performance of FC in VANETs is also studied in the literature. In [14], Maio et al. proposed a centralized approach to improve FC performance by optimizing the anchor zone size with the help of a centralized Software Defined Network (SDN) controller. They showed that having a centralized approach for obtaining key parameters like node density, mean speed etc., helps in fine tuning the core FC parameters like anchor zone radius. Optimizing the FC parameters achieves the desired application level performance. Moreover, it also maximizes the FC efficiency by minimizing the cost in terms of storage and bandwidth resources. In [15] we presented an analytical model for performance analysis of FC in vehicular environment. This model is based on a modified version of random waypoint mobility model and the novel thing about this model is that its derivation is not based on any specific road grid geometry. By doing extensive simulations under both synthetic and real-world vehicular traces, we proved that the proposed model accurately predicts the performance of FC for different applications. It is also shown that FC is feasible in realistic urban setups in almost all traffic conditions.

From the experimental perspective of FC, a few works exists in the literature. In [16] we presented the first experimental evaluation of FC in an office setting. The results from the experiment proved that under mobility patterns in an indoor office setting, FC is capable of providing very good performance in terms of success probability and availability. In [5], we presented the first experimental study of the performance of FC in a university campus context. During the experiment, the mobility patterns unveiled the key relevance of group dynamics in the user movements for the FC performance. In contradiction with the existing models, it is observed that a relatively low density of users is enough to ensure content persistence over time. A simple and preliminary version of our *Poisson Jumps model* was also proposed in [5], which is capable of computing a few performance metrics in a campus environment. However, that model does not capture the effects of clustering and correlation among mobility patterns of users. The current paper fills this gap by broadly extending our previous analytical work [5] and by proposing a new analytical model capable of capturing the effects of clustering and other user mobility characteristics in a campus/large scale office environment.

9. Conclusions

In this paper, we have proposed novel models for the characterization of Floating Content applications. Specifically, we have derived closed-form expressions for content item availability and for the success probability with which a user entering the anchor zone can get an item. Our analysis is based on the observation that in many environments mobile users spend most of the time in stationary positions and form groups.

With the aid of the model, we have investigated the operation of Floating Content. Our study unveils the importance of an *encounter index* that (i) allows to evaluate whether the system meets the theoretical conditions under which content can float with high probability, (ii) is semantically equivalent to the commonly used availability metric, and (iii) is key to compute the success probability.

We have also used real experiments to validate the model in a university campus context. Notably, our experiments show that, even in a highly dynamic setting, a steady state analysis like ours is very valuable for system design purposes. Besides, the experiments have demonstrated that a relatively low user density is enough to guarantee content persistence over time when, differently from common assumptions, users form groups and move to and from *attractor points*.

More specifically, we have shown that using just one seeder, the special type of user movement in the environment that we considered makes the availability of generated content items grow fast in the first hour, reaching on average about 50% of the users in the AZ. Whether this is acceptable for FC services, depends on the specific application, but we observed that it is easy to reach much higher availability values by just increasing the number of seeders to two. Indeed, by doing so, the number of users reached after one hour becomes about 90% on average. This provides quite interesting indications for the viability and the implementation of FC services in university settings, and in general in office environments.

Appendix A. Proofs of results for the Poisson Jumps mobility model (Section 3)

Proof. (Lemma 1) Let us consider the case in which the two nodes cannot jump out of the AZ. Assume first that one node is fixed, and the other jumps. As the location after the jump is uniformly distributed in the AZ, the chance of jumping within the transmission range of the fixed node is the ratio between the coverage area of a node, πr^2 , and the area of the AZ, πR^2 . Then, the number of jumps needed for the jumping node to come in range of the fixed one is geometrically distributed, with mean $\frac{R^2}{r^2}$. As $1/\mu$ is the mean duration of the stopping time, the mean amount of time for the two nodes to meet is $\frac{R^2}{\mu r^2}$. The inverse of this quantity gives the frequency at which such an event takes place. If we consider now that both nodes jump, the frequency with which the two nodes meet is two times greater than the one obtained when one is fixed. \Box

Proof. (Lemma 2) With average arrival rate γ and average sojourn time $T = \frac{1}{\mu(1-p)}$, the average number of users in the system is $N = \frac{\gamma}{\mu(1-p)}$ (Little's result). Hence the criticality condition (2) can be rewritten as in (3). The availability in steady state is the ratio between the average number of nodes with content, n, and the average population of the AZ, namely N, so that $\mathcal{A} = \frac{n}{N}$. The number of nodes in the AZ and the number of nodes the with content item are usually described with stochastic processes. Here instead we focus on the temporal evolution of the mean of these two system parameters. Since we assume to be in a stationary regime for the population of nodes in the AZ, denoting with N the mean number of nodes in the AZ, we can model the evolution of the average number of nodes n(t) with the content item over time by the following differential equation:

$$\frac{dn(t)}{dt} = \nu n(t)(N - n(t)) - n(t)\mu(1 - p)$$
(A.1)

The first term on the right is derived as follows. Assuming the content item to be uniformly distributed among nodes, the probability that a node has the content item at time *t* is $\frac{n(t)}{N}$. When two nodes come in range of each other, the probability that the content item is replicated is $2\frac{n(t)}{N}\left(1-\frac{n(t)}{N}\right)$. In the AZ at time *t* there are $\frac{N(N-1)}{2}$ possible pairs of nodes. Thus, the average rate at which the content item is transferred in the AZ at time *t* is given by $v\frac{(N-1)n(t)(N-n(t))}{N} \cong vn(t)(N-n(t))$.

The expression of v is given by Lemma 1. The second term is derived considering that $N\mu$ gives the average jumping rate within the AZ. Since only (1 - p) of those jumps goes out of the AZ, and only a fraction $\frac{n(t)}{N}$ of the jumping nodes possesses the content, on average, the rate at which nodes with the content item jump out of the AZ is $n(t)\mu(1-p)$.

Finally, solving for the stationary case (i.e., for n(t) = n and $\frac{dn(t)}{dt} = 0$), we get $n = N - \frac{\mu(1-p)}{\nu}$. If the criticality condition $\eta > 1$ holds, then 0 < n < N, so that *n* is well defined. Indeed, in that case $\mathcal{A} = \frac{n}{N} = 1 - \frac{\mu(1-p)}{\nu N} = 1 - \frac{1}{\eta}$, which is (4), and *n* can be expressed as in (5). Note that, if $\eta = 1$, both \mathcal{A} and *n* are 0, and the content does not float. If $\eta < 1$, there is no positive solution for the balance equation for *n*, so that the content does not float either, although in that case expressions (4) and (5) cannot be used. \Box

Proof. (Theorem 1) We assume border effects are negligible, which is a reasonable approximation when $r \ll R$. The probability of getting the content item at the arrival at a new waypoint can be written as

$$p_{jump} = \sum_{j=0}^{\infty} \Pr(j \text{ neighbors})[1 - \Pr(0 \text{ out of } j \text{ neighbors have the item})]$$
(A.2)

where Pr(*j*neighbors) is the probability of having *j* nodes in an area equal to πr^2 within the AZ of area πR^2 . The distribution of neighbors is therefore a Poisson one with average $N\frac{r^2}{R^2}$. For the computation of Pr(0out of prighbors have the item), we consider that in a stationary regime, and given that the location of each jump is uniformly distributed in the AZ, the probability for a node to have the content item is n/N, independently of its position in the AZ. Hence we have the following result:

$$p_{jump} = \sum_{j=0}^{\infty} \frac{(Nr^2/R^2)^j}{j!} e^{-Nr^2/R^2} \left[1 - \left(1 - \frac{n}{N}\right)^j \right]$$

Then we have that

$$\sum_{j=0}^{\infty} \frac{(Nr^2/R^2)^j}{j!} e^{-Nr^2/R^2} = 1$$

So that

$$p_{jump} = 1 - \sum_{j=0}^{\infty} \frac{(Nr^2/R^2)^j}{j!} e^{-Nr^2/R^2} \left(1 - \frac{n}{N}\right)^j = 1 - e^{-nr^2/R^2}$$

Considering the definition of α , we get (7).

For the derivation of p_{static} , for the law of total probability, we have

$$p_{static} = \int_0^{+\infty} f_{\mu}(t) p_{static|t} dt$$

*p*_{staticlt} is the probability of getting the content during a stop time of duration *t*, from nodes jumping within transmission range. Again, by the law of total probability,

$$p_{static|t} = \sum_{j=1}^{+\infty} \Pr(j \text{ jumps in } t) \left[1 - \left(1 - \frac{n}{N} \right)^j \right]$$

where Pr(*j* jumps in *t*) is the probability if having *j* nodes jumping within range (and hence resulting into a contact) during a stop time of duration t. $[1 - (1 - n/N)^j]$ is the probability that at least one out of these *j* nodes has the content. Again, we have

$$Pr(j \text{ jumps in } t) = \sum_{k=j}^{\infty} Pr(j \text{ jumps in } t|k)P(k, t)$$

where $Pr(j \mid p \mid k)$ is conditioned to having k jumps taking place in the AZ during the time interval of duration t. P(k, t) is the probability of having k jumps over time t. P(k, t) is reasonably well approximated by a Poisson distribution, of intensity $\gamma/(1-p)$. The approximation is due to the correlation between the process of exogenous arrivals and the counting process of the number of jumps inside the AZ.

We use a Poisson distribution, so that formally we write $P(k, t) = \frac{1}{k!} \left(\frac{t\gamma}{1-p}\right)^k e^{-\frac{t\gamma}{1-p}}$.

Given that waypoints are uniformly distributed at random within the $AZ_{i} \alpha$ (that is, the ratio between the area within range of a given node, and the AZ area) is a measure of the probability for a jumping node to arrive within range of the given still node. Hence,

$$\Pr(j \text{ jumps in } t)|k) = \alpha^j (1-\alpha)^{k-j}$$

 $\pm \infty$

Putting all together, we get expression (8) for p_{static} . $(1 - p_{stop})^{k+1}$ is the probability of not getting the content item during a sojourn in the AZ in which a node jumped k + 1 times into the anchor zone (including the first jump with which the node entered the AZ). By summing over all possible numbers of jumps k of a user during its sojourn in the AZ, we get (6). \Box

Proof. (Theorem 2) The derivation of the expression of $P_s(\tau)$ follows the same procedure as the one for p_{static} in Theorem 1. As for $P_m(\tau)$, we compute it as

$$P_m(\tau) = \sum_{i=1}^{+\infty} \Pr(i \text{ jumps in } [0, \tau]) [1 - (1 - p_{jump})^i]$$

As the stop time duration is exponentially distributed with mean $1/\mu$,

$$P_m(\tau) = \sum_{i=1}^{+\infty} \frac{(\mu \tau)^i}{i!} e^{-\mu \tau} [1 - (1 - p_{jump})^i] = 1 - e^{-\mu \tau p_{jump}} \Box$$

Appendix B. Proofs of results for the Attractor-based mobility model (Section 4)

Proof. (Lemma 3) The proof goes as the one for Lemma 1, with the following differences. The frequency with which two nodes come in contact, is still $v = 2\mu\alpha$. However, when a node jumps within the AZ in presence of clustering, the probability of falling in range of the other node (probability which we have denoted with α) is not anymore equal to the ratio between the area in range of the node (πr^2) and the AZ area (πR^2), as clustering makes some locations of the AZ more likely to receive the jumping node than others. Since there are $\beta\pi R^2$ APs, the probability for the jumping node to fall on the same AP as the other node is $\frac{1}{\beta\pi R^2}$. Therefore, we can write

$$a = \frac{1}{\beta \pi R^2} + \left(1 - \frac{1}{\beta \pi R^2}\right) \frac{r^2}{R^2}$$

That is, if the node does not jumps to the same AP as the other, it may still fall on another AP which is within the transmission range of that node. As APs are uniformly distributed, this probability is still r^2/R^2 . \Box

Proof. (Lemma 4) Let us consider an AZ with *B* APs. Every time a node enters the AZ, the probability to land at a given AP is $\frac{1}{B}$. In stationary conditions, the probability for an AP to host at least one node is therefore

$$1 - \left(1 - \frac{1}{B}\right)^N$$

and hence

dm(t)

$$M = B \left[1 - \left(1 - \frac{1}{B} \right)^N \right] \tag{B.1}$$

Now, following the same line of reasoning as in Lemma 2, we have the following differential balance equation for the number of APs with at least one node with content:

$$\frac{dm(t)}{dt} = \nu m(t)(M - m(t))K - Hm(t)\mu(1 - p)$$
(B.2)

where $K = \frac{N}{M}$ is the average size of a cluster gathered at an AP. *H* is the fraction of active APs which have *exactly* one node, which is equivalent to the probability that an AP has only one node:

$$H = {\binom{B}{1}} \frac{1}{B} \left(1 - \frac{1}{B}\right)^{N-1} = \left(1 - \frac{1}{B}\right)^{N-1} \stackrel{=}{=} \frac{1 - \frac{M}{B}}{1 - \frac{1}{B}}$$
(B.3)

where the last passage comes from solving (B.1) for $(1 - \frac{1}{B})^{N-1}$. Now, by replacing *B* with its average $\beta \pi R^2$ in (B.1) and in (B.3), we obtain (15) and (14), respectively. Furthermore, in stationary conditions, setting $\frac{dn(t)}{dt} = 0$ and $m = \lim_{t\to\infty} m(t)$ in the (B.2) yields

$$m = M - \frac{H\mu(1-p)}{K^2\nu} = M\left(1 - \frac{H}{K}\frac{\gamma}{N^2\nu}\right)$$
(B.4)

which must be a proper value 0 < m < M only if the criticality condition $\eta > 1$ holds. Therefore, the criticality condition for the AB model becomes (12). Note also that if $\eta \le 1$ in (12) the balance equation has no positive solution, which means that there are no APs with nodes with content items in steady state (i.e., the content does not float). For what concerns the expression of the availability, consider that N/M is the average number of nodes per active AP, and $m \cdot \frac{N}{M}$ is therefore the total number of nodes with content, which is n by definition. Thus, m/M = n/N, and hence (13). Finally, substituting the expressions of K and H in (B.4), we have (16).

Proof. (Theorem 3) Again, we assume border effects are negligible and consider the presence of N nodes and B APs. Let us denote with X the probability of having none of the nodes at a given AP within the AZ:

$$X = \left(1 - \frac{1}{B}\right)^{N}$$

Since an AP is active with probability 1 - X, then the number of active APs in the AZ is M = (1 - X)B and hence the probability that an AP is active and has content is $(1 - X)\frac{m}{M} = \frac{m}{B}$.

Note that a jumping node is always at least in range of the cluster at its destination waypoint. Hence, the probability of getting the content item at the arrival at a new waypoint is

$$p_{jump} = \Pr(\text{dest waypoint}) + (1 - \Pr(\text{dest waypoint}))$$

$$\cdot \sum_{j=0}^{\infty} \Pr(j \text{ active APs in range}|\text{waypoint})[1 - \Pr(0 \text{ out of } j \text{ active APs have the content})]$$
(B.5)

where $Pr(\text{dest waypoint}) = \frac{m}{B}$ because it is the probability that the destination waypoint is an active AP with nodes with content, and $B = \beta \pi R^2$ to account for the entire AZ. Moreover, Pr(jactive APs in range|waypoint) is the probability of having in range *j* active APs, *in addition to the destination AP*. Being the point process of *active* APs a marked Poisson process derived from the point process of APs within

 $\Pr(j \text{ active APs in range}|\text{waypoint}) = \frac{\left[(1-X)\beta\pi r^2\right]^j}{j!}e^{-(1-X)\beta\pi r^2}$

Eventually, the probability that an active AP has content is, by definition, $\frac{m}{M}$. Thus, putting it all together, (B.5) becomes

$$p_{jump} = \frac{m}{\beta \pi R^2} + \left(1 - \frac{m}{\beta \pi R^2}\right) \sum_{j=0}^{\infty} \frac{\left\lfloor (1 - X)\beta \pi r^2 \right\rfloor^j}{j!} e^{-(1 - X)\beta \pi r^2} \left[1 - \left(1 - \frac{m}{M}\right)^j\right] \\ = \frac{m}{\beta \pi R^2} + \left(1 - \frac{m}{\beta \pi R^2}\right) \left[1 - e^{-(1 - X)\beta \pi r^2 \frac{m}{M}}\right]$$
(B.6)

With the expression for *M* found at the beginning of the proof (with *B* replaced by its average value $\beta \pi R^2$) the above expression (B.6) leads to (17).

The derivation of p_{static} is the same as for Theorem 1. The only difference here is in the expression of α which is a measure of the probability for a jumping node to arrive within range of the given still node. In the AB model, the expression for α is given by Lemma 3.

References

- [1] J. Alcaraz-Calero, I.P. Belikaidis, C.J.B. Cano, P. Bisson, D. Bourse, M. Bredel, D. Camps-Mur, T. Chen, X. Costa-Perez, P. Demestichas, M. Doll, S.E. Elayoubi, A. Georgakopoulos, A. Mammela, H.P. Mayer, M. Payaro, B. Sayadi, M.S. Siddiqui, M. Tercero, Q. Wang, Leading innovations towards 5G: Europe's perspective in 5G infrastructure public-private partnership (5G-PPP), in: Proceedings of the 2017 PIMRC, 2017, pp. 1–5. doi:10.1109/PIMRC.2017.8292654.
- [2] A. Asadi, Q. Wang, V. Mancuso, A survey on device-to-device communication in cellular networks, IEEE Commun. Surv. Tutor. 16 (4) (2014) 1801–1819. doi: 10.1109/ COMST.2014.2319555
- [3] E. Hyytiä, J. Virtamo, P. Lassila, J. Kangasharju, J. Ott, When does content float? Characterizing availability of anchored information in opportunistic content sharing, in: Proceedings of the 2011 INFOCOM, Shanghai, China, 2011, pp. 3123–3131.
- [4] S. Ali, G. Rizzo, B. Rengarajan, M. Ajmone Marsan, A simple approximate analysis of floating content for context-aware applications, in: Proceedings of the 2013 MobiHoc, 2013, pp. 271–276.
- [5] S. Ali, G. Rizzo, V. Mancuso, M. Ajmone Marsan, Persistence and availability of floating content in a campus environment, in: Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM), IEEE, 2015, pp. 2326–2334.
- [6] T. Spyropoulos, K. Psounis, C.S. Raghavendra, Performance analysis of mobility-assisted routing, in: Proceedings of the 2016 MobiHoc, 2006, pp. 49–60.
- [7] M.S. Desta, E. Hyytia, J. Ott, J. Kangasharju, Characterizing content sharing properties for mobile users in open city squares, in: Proceedings of the 2013 WONS, 2013, pp. 147–154. doi:10.1109/WONS.2013.6578340.
- [8] P. Serrano, A. De La Oliva, P. Patras, V. Mancuso, A. Banchs, Greening wireless communications: status and future directions, Comput. Commun. 35 (14) (2012) 1651–1661.
 [9] E. Hyytia, P. Lassila, J. Ott, J. Kangasharju, Floating information with stationary nodes, in: Proceedings of the 2012 WiOpt, 2012, pp. 361–366.
- [10] A.A.V. Castro, G. Di Marzo Serugendo, D. Konstantas, Hovering information self-organising information that finds its own storage, in: Proceedings of the 2008 SUTC, 2008, pp. 193–200.
- [11] J. Ott, E. Hyytia, P. Lassila, T. Vaegs, J. Kangasharju, Floating content: information sharing in urban areas, in: Proceedings of the 2011 PerCom, 2011, pp. 136–146.
- [12] S. Ali, G. Rizzo, M. Ajmone Marsan, V. Mancuso, Impact of mobility on the performance of context-aware applications using floating content, in: Proceedings of the 2013 ICCASA, 2013.
- [13] R. Hagihara, Y. Yamasaki, H. Ohsaki, On delivery control for floating contents sharing with epidemic broadcasting, in: Proceedings of the 2017 CCNC, 2017, pp. 353–356. doi:10.1109/CCNC.2017.7983133.
- [14] A.D. Maio, R. Soua, M.R. Palattella, T. Engel, G.A. Rizzo, A centralized approach for setting floating content parameters in VANETs, in: Proceedings of the 2017 CCNC, 2017, pp. 712–715. doi:10.1109/CCNC.2017.7983220.
- [15] G. Manzo, M. Ajmone Marsan, G. Rizzo, Performance modeling of vehicular floating content in urban settings, in: Proceedings of the Twenty-Ninth International Teletraffic Congress (ITC 29), 1, 2017, pp. 99–107. doi:10.23919/ITC.2017.8064344.
- [16] S. Ali, G. Rizzo, V. Mancuso, V. Cozzolino, M. Ajmone Marsan, Experimenting with floating content in an office setting, IEEE Commun. Mag. 52 (6) (2014) 49–54 doi:10.1109/MCOM.2014.6829944.



Gianluca Rizzo is Senior Research associate at HES-SO since April 2013. Previously, he has been Research Assistant Professor at IMDEA Network, and Adjunct Professor at UC3M. He has been with the IC lab at EPFL, Switzerland, between 2004 and 2008, where he received his Ph.D. in Computer Science in 2008. His main research interests are in the design, modeling and performance evaluation of distributed systems. He has been working on design and evaluation of algorithms for dynamic management of wireless networks, on opportunistic communication paradigms such as Floating Content, on QoS provisioning in vehicular networks, on issues of IoT integration, and on QoS guarantees in avionic networks.



Vincenzo Mancuso is Research Associate Professor at IMDEA Networks Institute, Madrid, Spain, and recipient of a Ramon y Cajal research grant of the Spanish Ministry of Science and Innovation. Previously, he was with INRIA Sophia Antipolis (France), Rice University (Houston, TX, USA) and University of Palermo (Italy), from where he obtained his MSc and his Ph.D. in Electronics, Computer Science and Telecommunications. He has authored more than 100 peer-reviewed publications focusing on Internet QoS and on the analysis, design, and experimental evaluation of opportunistic and adaptive protocols and architectures for wireless networks. He is currently working on analysis and optimization of wireless networks, on the analysis and design of IoT solutions, and on the measurements and assessment of mobile broadband networks.



Shahzad Ali received his M.Sc. in Telematics Engineering in 2011 and his Ph.D. in Telematics Engineering from University Carlos III of Madrid, Spain in 2014. Currently, he is working as assistant professor at department of Computer Science, Jouf University, Tabarjal, Saudi Arabia. His research interests include performance analysis of context-aware applications, wireless sensor networks, vehicular ad hoc networks, and opportunistic networks. He has published his research work in top conferences like INFOCOM, Mobinoc, etc. and well-reputed journals of his field. He is also acting as a reviewer of many well-reputed journals and also acting as TPC for many well-known conferences.



Marco Ajmone Marsan is a full professor at the Electronics and Telecommunications Department of the Politecnico di Torino in Italy, and a part-time research professor at IMDEA Networks Institute in Leganes, Spain. He obtained degrees in EE from the Politecnico di Torino and the University of California, Los Angeles (UCLA). Since 1974 he has been at Politecnico di Torino, in the different roles of an academic career, with an interruption from 1987 to 1990, when he was a full professor at the Computer Science Department of the University of Milan. He has been doing research in the fields of digital transmission, distributed systems and networking. He has published over 350 papers in the leading conferences and journals of his research area as well as two books. Marco Ajmone Marsan has been a member of the editorial board and of the steering committee of the "ACM/IEEE Transactions on Networking". He is a member of the editorial boards of the journals "Computer Networks" and "Performance Evaluation" of Elsevier and of ACM TOMPECS. He was in the organizing committee of several leading networking conferences, and general chair of INFOCOM 2013. He is a Fellow of the IEEE, and a member of the Academia Europaea and of the Academy of Sciences of Torino. He received a honorary degree in Telecommunication Networks from the Budapest University of Technology and Economics. He was the Vice-Rector for Research, Innovation and Technology Transfer at the Politecnico di Torino, and the Director of the IST the Just della Informazione e delle Telecomunicazioni of the Italian National Research Council. He was the Italian delegate in the ICT and IDEAS committees of FP7.