A Walk Down Memory Lane: On Storage Capacity in Opportunistic Content Sharing Systems

Gianluca Rizzo,[1] Noelia Perez Palma,[2,3] Marco Ajmone Marsan,[2,4] and Vincenzo Mancuso[2]
gianluca.rizzo@hevs.ch, noelia.perez@imdea.org, ajmone@polito.it, vincenzo.mancuso@imdea.org
[1] HES SO Valais [2] Institute IMDEA Networks [3] University Carlos III of Madrid [4] Politecnico di Torino

***Abstract*—Floating Content (FC) is a paradigmatic example of opportunistic infrastructure-less content sharing system where information is spread upon mobile node encounters within an area which is called the replication zone. FC allows the probabilistic spatial storage of information, even in the case of unreliable communications, with no support from dedicated servers. Given the large amount of communication and storage resources typically required to guarantee content persistence despite node mobility, a major open issue for the practical viability of FC and of similar distributed storage systems is the characterization of their storage capacity, i.e., of the maximum amount of information which can be stored for a given set of system parameters. In this paper, we propose a simple yet powerful information theoretical model of the storage capacity of probabilistic distributed storage systems such as FC, based on a mean field model of opportunistic information exchange. We evaluate numerically our results, and validate the model by means of realistic simulations, showing the accuracy of our mean field approach and characterizing the properties of the FC storage capacity versus the main system parameters.**

I. Introduction

In light of the rapid growth as well as the general underutilization of storage capacity in mobile devices, distributed storage schemes have been proposed to reduce wireless capacity bottlenecks by caching popular content items in mobile UEs (i.e., at the extreme edge of the network) [1]–[3]. By enabling direct content exchange between mobile users without routing through base stations (BSs), distributed edge storage systems can achieve higher energy efficiency while decreasing the congestion of BS resources. Among these techniques, an important role is played by those opportunistic communication paradigms for probabilistic dissemination of contextualized information, denoted as Floating Content [4], or Hovering Information [5]. These approaches (which we henceforth denote as *FC*), all aim at the local dissemination of information to end users over a defined geographic area (called replication zone or RZ), based solely on direct device-to-device (D2D) connectivity [6]. Through opportunistic replications, they store information spatially, in a probabilistic fashion, despite the mobility of UEs and the unreliability of information exchanges, with no need for centralized servers. Their goal is to deliver the stored content proactively to those users which are expected to traverse a specific region (the Zone of Interest, or ZOI), before they enter the region, while minimizing support from the cellular infrastructure. Hence, the main performance metric in such systems is the *success ratio*, i.e., the average fraction of nodes that enter the ZOI with content over a given time interval.

Clearly, guaranteeing (probabilistically) content persistence and a given target performance in such a volatile setting without the support of a centralized, static infrastructure comes at the cost of a drastic increase both in content redundancy across the user population, and in the communications required in order to serve the target population of users with respect to classic, infrastructure-based solutions. As a consequence, a major open issue for the practical viability of such distributed edge storage systems is the characterization of their capacity, i.e., of the amount of information they are able to store for a given set of system parameters. Such characterization is key for D2D resource allocation, and for the design of incentive schemes for cooperation, as it enables the estimation of the marginal contribution of each resource to the overall service capacity, and hence the appropriate tuning of resource costs and the proper setting of rewards for cooperation.

Given the infrastructure-less nature of these paradigms, several performance studies have focused on the conditions under which content *floats*, i.e., persists in such a distributed storage scenario. [4] introduces the *criticality condition*, a sufficient condition for the content to float indefinitely with very high probability, under various mobility models. [7] introduces an analytical model to estimate content persistence in the case of outdoor pedestrian mobility over large open spaces, such as a city square. [8], [9] characterize the mean time to information loss, on several scenarios, based on synthetic mobility and on measurement-based vehicular mobility traces. [10] further explores the capabilities of FC in urban mobility scenarios, for which the authors illustrate an effective modeling framework. Other works, e.g., [11], [12], focus on how to engineer the replication and storage strategies in realistic settings in order to efficiently guarantee a given success probability within a predefined temporal interval. [13] discusses how to adapt the FC geographical scope according to the context, so to optimize performance. However, all these approaches only consider scenarios with a single content item, ignoring those issues arising when several different contents share the same set of UEs for replication and storage.

We propose a simple analytical model of the storage capacity of probabilistic distributed edge storage systems such as FC, based on a mean field model of the opportunistic information exchange, which allows for a first order characterization of the scaling laws of the storage capacity of these systems. Specifically, the contributions of this paper are as follows:

- We develop an analytical model of FC performance, based on a mean field description of the dynamics of the population of users storing the information items, and of the population of users which are in the process of exchanging (sending or receiving) such items;
- We derive the expressions of the storage capacity, as a function of node mobility and of the geometry of the replication zone. We formulate an optimization problem for the derivation of the maximum amount of information which can be stored with FC, showing that it can be solved efficiently. To the best of our knowledge, this is the first work to characterize analytically the storage capacity of probabilistic distributed storage schemes such as FC;
- We evaluate numerically our results, validating our assumptions against simulation, showing the accuracy of our mean field approach, and characterizing the properties of FC storage capacity as a function of the main system parameters.

The paper is organized as follows. In Section II, we present the system model, and in Section III we present our mean field approach to performance analysis of FC. In Section IV we characterize the storage capacity of FC. In Section V we numerically assess the accuracy of our results, and evaluate the impact of the main system parameters on FC storage capacity. Finally, Section VI concludes the paper.

II. System Model

We consider a region of the plane where nodes move according to a stationary mobility model such that, at any time instant, nodes are uniformly distributed in space (however, the impact of a non-uniform node distribution will be assessed by simulation). Each node is equipped with a wireless transceiver. We assume each node knows its position in space. We say two nodes are *in contact* when they are able to directly exchange information via wireless communications. With τ we denote the contact time, i.e., the duration of the time interval during which two nodes are in contact, and with $f(\tau)$ we indicate its PDF. With C_0 we denote the mean channel capacity between two nodes that are in contact. Let D be the mean node density in the given region, and g the mean rate of contacts per unit area. We focus on those settings in which mean node density and transmission power are such that store-carry-and-forward is the main mode of content diffusion and communication. Indeed, when this is not the case, other, more efficient techniques than FC are applicable for content diffusion.

Let x denote a position in space and θ an angle, and let $\alpha(x, \theta)$ be the *angular node flow*, i.e., the rate of nodes moving in direction $(\theta, \theta + d\theta)$ across a small perpendicular line segment of length ds centered at x, divided by $ds \cdot d\theta$. Let α denote the *linear flow*, given by

$$\alpha = \int_0^{\pi} \alpha(x, \theta) d\theta. \tag{1}$$

α is the rate at which nodes traverse a segment of unit length centered at x in a same direction (e.g., from left to right of the segment). In order to simplify notation, in what follows we assume that the mobility model is isotropic, and hence, that α does not depend on the position in space of the unit segment, nor on the specific direction of the node flow. However, note that our approach can be easily extended to more general, non-isotropic mobility models.

TABLE I: Main notation used in the paper.

Notation	Parameter
M	Data storage capacity of each node ($bits$)
D	Node density (m^{-2})
C_0	Mean capacity of the link between two nodes in contact ($bits/s$)
L	Content size ($bits$)
R	Replication Zone (RZ) radius (m)
g	Mean contact rate per unit area ($s^{-1}m^{-2}$)
α	Linear flow ($s^{-1}m^{-1}$)
τ_0	Transfer setup time (s)
τ	Contact time (s)
$f(\tau)$	PDF of contact time

A. Floating Content basic operation

We briefly recall the basic operation of the Floating Content (FC) communication paradigm [4]. At time t_0, a node (the *seeder*) generates a piece of content (e.g., a text message, or a picture). The *replication zone* (RZ) is the region of the plane (containing the location of the seeder at t_0) in which, for $t \geq t_0$, when a node with content comes in contact with a node without it (both located in the RZ), content is exchanged. When a node leaves the RZ, we assume it discards the content.

We assume that before each content transfer, a setup time τ_0 is required, which models the time taken by each node to detect the presence of the other node, to verify the availability of content and to set up the transfer. With S we denote the probability that a content transmission completes successfully. This happens when the contact time and the channel capacity between the two nodes during contact time are such that content can be transferred in full. We assume that every node can exchange content (i.e., send or receive) with one node at a time, though other modes of communications, e.g., broadcast, can be easily accounted for in our approach.

The goal of the FC paradigm is typically to ensure, through opportunistic replication, that the content item is delivered to a given mean percentage of those nodes entering a predefined limited geographical region within the RZ (the *zone of interest* or ZOI), by the time they enter in such region. One of the main system parameters for FC is content *availability* at time $t \geq t_0$, i.e., the mean fraction of nodes with content, at time t in the RZ. High values of availability in the RZ typically correlate with low likelihood of content disappearance, and with high probability of transferring content to nodes entering the RZ. Given a time interval during which the content is floating in the RZ, the FC *success ratio* P_{succ} is the mean fraction of users which enters the ZOI with content during such time interval. The target value of the success ratio, as well as the choice of size, shape and location of the ZOI are functions of the performance requirements of the specific application and service supported by FC. For instance, in an application whose performance target is to deliver an advertisement (e.g., info on a sale) to at least a given percentage of the people who enter a shopping center, the borders of the ZOI should include all entrances to the shopping center.

From this description, it emerges clearly that in such an opportunistic communications scheme, when proper conditions (in terms of user density and mobility, and of size of the RZ) are met, the content *floats* (i.e., it persists probabilistically) in the RZ, even after the seeder has left it. In practice, the content never fluctuates forever, as the likelihood of content disappearance cannot be brought to zero. We denote the system corresponding to one content item floating in its RZ according to the scheme described above as a *Floating Element* (FE). In this work, we focus on systems generally composed by several floating elements, in which for simplicity we assume that all RZs are circular with radius R, and each ZOI is strictly enclosed in its RZ. Moreover, all contents have the same size L, and M is the size of memory that can be used for FC in each node. Note however that our approach can be extended to contents of different size and to scenarios where nodes have different memory sizes, mainly at the cost of increasing the notation complexity.

With respect to the relative position of the RZ of each floating element, we consider two types of systems. In *distributed floating systems*, the centers of the RZ of each floating element are distributed according to a Poisson Point Process (PPP) with intensity γ. This scenario accounts for the partial RZ overlap which is typical of setups where each application or end user defines its own ZOI location. In *localized floating systems*, we assume all FEs share a same RZ and ZOI. In this case, the *intensity* γ is the ratio of the total number of FEs, over the RZ area. This second scenario allows ruling out the impact of randomness in RZ overlapping on the storage capacity of the system, and hence it allows investigating the maximum amount of information which can be stored in a given location (the RZ) via the FC paradigm.

We assume that at each contact, every time two nodes have to decide which content to transfer, each content which could be transferred (because it is stored in one node but missing in the other, and because the node missing it has enough memory space to store it) has the same probability to be chosen for the transfer, regardless of which of the two hosts it resides on. We also assume content exchanges between two nodes to be unidirectional. These assumptions do not limit the validity of our approach, which can easily be expanded to include, e.g., bidirectional content exchanges, the effects of content broadcasting, or any specific priority scheme for contents to be exchanged.

III. A Mean Field Model of FC

In FC, the set of nodes exchanging content within a RZ can be modeled as a system of interacting objects, in which interactions bring to changes in the state of the objects via content replication. When the number of objects becomes large, the analytical performance study of such system becomes difficult, due to the exponential growth of the state space size.

To define the system state, at any time $t \geq 0$, we associate with each node two binary vectors. In the first vector, the j-th element is equal to one if the j-th content is stored in the node, and zero otherwise. In the second vector, the j-th element indicates whether the node is inside the RZ for the j-th content. In addition, with each node we associate another state variable, which is equal to one if the node is busy exchanging (transferring or receiving) contents with another node, and zero otherwise. Defining as system state the collection of all node variables, we clearly see that the number of states explodes as γ (and therefore the number of different contents) and the number of nodes grow. Such a system can be assumed to evolve according to Markovian dynamics, and it can therefore be modelled as a Continuous Time Markov Chain (CTMC).

In order to derive meaningful insights into the performance of such systems, in what follows we adopt a technique based on the *mean field interaction model* or *fluid limit* [14], hence on an approximate model of the interactions between nodes. The first approximation step of such approach is based on assuming that the following *homogeneous conditions* hold:

Definition 1 (**Homogeneous conditions**)**.** We say that a floating system satisfies the homogeneous conditions if:

- at $t = 0$ the mean number of nodes per unit surface possessing a given content is the same for all contents;
- at any time instant, nodes possessing a given content are uniformly distributed within the RZ for that content; and
- the probability of a node to have that content is independent from the probability of any other node to have the same content.

The homogeneous conditions assumption (equivalent to, e.g. the "well stirred" assumption in chemistry [15], and to the assumption of stochastic equivalence of nodes within a same class in [16]) allows deriving simpler expressions for the evolution of the main performance parameters of the system, at the cost of neglecting spatial inhomogeneities. Note however that our approach can be extended to account for spatial variations as well as for possible nonuniform seeding strategies (e.g., through the notion of node class, as in [16]), though at the cost of an increase in complexity of the analytic expressions.

In order to model the temporal evolution of content diffusion and availability in our system, we focus on the temporal evolution of two classes of node populations. The first class is composed by those nodes possessing a given content at a given time instant, while the second is the set of those nodes which are *busy*, i.e., exchanging contents, at a given time. Note that the busy state is not associated with a given content, but with the fact that the node is involved in an exchange of contents at a given time instant, and that each node can be part of both classes of populations at the same time.

As a consequence of the homogeneous conditions, and of the random scheduling of content transfers, at any time instant the PDF of the fraction of nodes possessing the j-th content within the RZ for that content (i.e., the PDF of the *availability* for content j) is the same for all contents j, in both distributed and localized floating systems. Hence, in what follows, for a given choice of RZ radius R, we indicate with $a(t,R)$ the mean availability at time t over all contents. The following result derives the PDF of the number of contents possessed by a node.

Lemma 1. *With the given assumptions, in a floating system (distributed or localized) the number of contents possessed by a node at time t, denoted as $m(t,R)$, is distributed as a binomial $Bin(n,p)$, with parameters $n = \lfloor \gamma\pi R^2 \rfloor$ and $p = a(t,R)$, and truncated in $\lfloor \frac{M}{L} \rfloor$.*

Proof: In distributed floating systems, the number of contents which a node can possess and replicate at a given point in time is equal to the minimum between the number of RZs in which the node is at that time, whose mean is $\gamma\pi R^2$, and the number of contents which can be stored in its memory. In localized floating systems, the average number of RZs in which a node is located is $\gamma\pi R^2$. Since the probability of possessing a given content is well approximated by $a(t,R)$, and since by the homogeneous assumption the probability for a node to possess a content is independent from other nodes, and is the same for all nodes, the number of contents possessed by a node follows a binomial distribution, truncated at $\lfloor \frac{M}{L} \rfloor$. ■

On the occurrence of a contact event, a relevant parameter is the mean amount of *exchangeable* contents, i.e., of contents which are possessed only by one of the two nodes, and for which there is enough free memory to store them at the receiving node. This parameter is key in determining the likelihood of a content to be exchanged on a contact event.

Lemma 2. *With the given assumptions, in a floating system (distributed or localized) the mean number of exchangeable contents on a contact at time t is given by $c(t,R) = 2E[x(t,R)]$, where $x(t,R)$ is a discrete random variable whose PDF is given by*

$$P(x(t,R)) = E_{m_a,m_b}\left[\frac{\binom{m_a}{x(t,R)}(1-a(t,R))^{x(t,R)}a(t,R)^{m_a-x(t,R)}}{F_{c_a}(u)}\right], \quad (2)$$

with $u = \min(m_a, \lfloor \frac{M}{L} \rfloor - m_b)$ and $F_{c_a}(u)$ is given by

$$F_{c_a}(u) = \sum_{k=0}^{u} \binom{m_a(t,R)}{k}(1-a(t,R))^k a(t,R)^{m_b(t,R)-k}.$$

m_a and m_b are the amount of contents possessed by the two nodes in contact at time t, whose distribution is given by Lemma 1. E_{m_a,m_b} denotes the expectation with respect to m_a and m_b.

For the proof, please refer to Section A in the Appendix.

Let $c(t,R)_\infty$ denote the mean number of exchangeable contents at a contact as per Lemma 2, when host memory is infinite. The probability of successful transfer of a *single* content during a contact of duration τ is a function of the ratio between τ and the amount of time required to transfer all exchangeable contents. Averaging over contact duration, we get

$$S(t,R) = \frac{c(t,R)}{c(t,R)_\infty}\int_{\tau_0}^{+\infty} \min\left(1, \left\lfloor \frac{(\tau-\tau_0)}{\frac{L}{C_0}} \right\rfloor \frac{1}{c(t,R)}\right) f(\tau)d\tau. \quad (3)$$

Similarly, the mean duration of an exchange of contents between two nodes (i.e., mean duration of the time interval during which a node is in the busy state) is given by

$$T_s(t,R) = \int_0^{+\infty} \min\left(\tau, \frac{c(t,R)L}{C_0} + \tau_0\right) f(\tau)d\tau. \quad (4)$$

Let us denote with $b(t,R)$ the fraction of busy nodes at time t for a RZ radius R. This parameter is key in order to model the decrease of the rate at which contents get replicated successfully when the mean time required to transfer the exchangeable contents is comparable or larger than the mean contact time. The following result models the asymptotic dynamics of our system over finite time intervals, for large RZ areas and hence for a large amount of nodes involved in the process of diffusion of each content.

Theorem 1. *In a floating system (distributed or localized), for any initial condition $(a(0,R), b(0,R))$ with $a(0,R) > 0$, $b(0,R) = 0$, for large R, the quantities $(a(t,R), b(t,R))$ converge almost surely over any finite horizon to the solution $(a(t), b(t))$ (the mean field limit) of the following ordinary differential equations (ODEs):*

$$\begin{cases} \dfrac{da(t)}{dt} = \dfrac{b(t)}{T_s(t)}a(t)(1-a(t))S(t) - \dfrac{2\alpha}{DR}a(t), \\ \dfrac{db(t)}{dt} = \dfrac{g}{D}2(1-b(t))^2 - \dfrac{b(t)}{T_s(t)} - \dfrac{4\alpha}{DR}b(t), \end{cases} \quad (5)$$

with initial condition $(a(0), b(0)) = (a(0,R), b(0,R))$. $S(t)$ and $T_s(t)$ denote the expressions in (3) *and* (4) *respectively, with $a(t)$ instead of $a(t,R)$.*

For the proof, please refer to Section B in the Appendix. This result states that the probability of observing a difference between any point of the trajectory of the given system and the solution of the ODEs goes to zero as R (and hence the mean total number of nodes in each RZ) grows. That is, in the limit, the error made by considering a deterministic system characterized by $a(t)$ and $b(t)$ instead of the actual system goes to zero. Moreover, at any time t, for any R, $a(t)$ is the expected value of $a(t,R)$. Note that we consider $b(0,R) = 0$ as we assume that at $t = 0$ no node has initiated any content transfer yet.

A. Quasi-stationary regime

In this work, we are mainly interested in the performance of the system after *enough time* has passed from the initial seeding of the content, i.e at times in which the dynamics of initial content diffusion are exhausted. However, in a finite system there is always a non-zero chance of having a content item disappear from its RZ due to random fluctuations in the population of nodes which possess it. The possibility of such an event is present in any finite floating system, which therefore for $t \to \infty$ inevitably tends towards the empty state. As the CTMC of the system has an absorbing state (the empty state), its only equilibrium state is the empty system.

Therefore, in what follows we focus on the performance of the system in its *quasi-stationary* regime, i.e., at a time from content seeding which is *large enough* for the initial dynamics of seeding to be exhausted, and at the same time *small enough* for content not to have been absorbed yet. Indeed, this is the performance regime which is the most relevant, as in any practical setting the amount of time during which a content should be stored and made available is not infinite (e.g., due to day/night patterns in vehicles and pedestrians mobility) but it is often long enough for the initial transient to have only a marginal impact on the overall performance.

The computation of an estimate of the time to content extinction based on the original CTMC is unfeasible due to the explosion in the number of the states which should be considered. In this section, under some mild assumptions, we derive a condition for the content to float for an indefinitely long amount of time (i.e., for the quasi stationary regime to exist), as well as expressions for the main performance parameters at times from content seeding in which the initial transient effects no longer influence the system behavior. In the numerical section we will assess the accuracy of our model, and we will investigate the conditions under which the decay of the system towards the empty state has a significant impact on system performance.

Given the existence of the absorbing state for the original system (which we denote as $\mathcal{S}$), in order to apply the mean field approach and compute the performance parameters for the time before absorption we consider a slightly different system (denoted with $\mathcal{S}'$). This new system is obtained from the original one by assuming that for each content, when the system is empty, content is re-seeded. Specifically, we assume that, when a content is absorbed in $\mathcal{S}'$, at rate ϵ (constant and independent on any parameter of the system) the process selects a node at random in the RZ, and injects the content in that node. Note that, in the time period between content seeding and content absorption, $\mathcal{S}$ and $\mathcal{S}'$ are indistinguishable. This is confirmed by the following result, relative to the mean field limit over finite time intervals:

Theorem 2. *For any $\epsilon \geq 0$ the mean field limit of system $\mathcal{S}$ and $\mathcal{S}'$ is the same.*

Proof: This result is due to the fact that the rate ϵ, being constant with RZ radius R, vanishes with increasing RZ radius at a rate R^{-2}, and is hence negligible in the mean field approximation, in which all neglected terms are $\mathcal{O}(R^{-2})$. ∎

Theorem 2 implies that the content seeding process has a vanishing impact on system performance in the mean field regime, as modeled in the previous section. As a result, the mean field approximation results of Theorem 1 hold also for $\mathcal{S}'$. This result allows exploiting Theorem 1 in order to derive a mean field approximation for the stationary state, which is only defined for $\mathcal{S}'$. Then, thanks to the fact that the two systems are indistinguishable at times between content generation and absorption, in a regime in which content absorption is relatively infrequent (i.e., in a regime when quasi-stationary state is defined) the mean field approximation for the stationary state of $\mathcal{S}'$ is also an approximation of the quasi-stationary regime of $\mathcal{S}$. In Section V we assess the accuracy of our approach.

The following result establishes a relation between the stationary state of $\mathcal{S}'$ and the steady-state solutions of the problem in (5).

Theorem 3. *For any $\epsilon > 0$, for large R the steady-state solutions of Equation (5) are an approximation of the state distribution of $\mathcal{S}'$ for $t \to \infty$.*

Proof: The CTMC associated to $\mathcal{S}'$ has no absorbing states, and its state diagram presents no cycles. Hence it belongs to the class of reversible stochastic processes [17]. Therefore, from Theorem 1.2 in [18] it follows that its stationary behavior is completely determined by the solutions of the ODEs (5). ∎

Let (a, b) denote the steady-state solutions of (5). The following result gives their expressions, as a function of the main system parameters.

Theorem 4. *For the considered system, the steady-state solution of ODEs (5) is given by the solution of the following fixed point problem in a:*

$$a = 1 - \frac{2\alpha T_s(a)}{DRS(a)b} \tag{6}$$

with $b = K - \sqrt{K^2 - 1}$, and $K = 1 + \frac{D}{4gT_s(a)} + \frac{\alpha}{gR}$. If the condition

$$\frac{2\alpha T_s(a)}{DRS(a)b} \leq 1 \tag{7}$$

is satisfied, and the system starts from a non-empty state (i.e., if at $t = 0$ there is at least one node with content for each content j), any trajectory of the system converges to the solution of Equation (6). Otherwise, the system converges to the empty state for any initial condition.

Proof: (sketch) Both the unicity of the fixed point solution, and the convergence derive from the study of the gradient of Equation (6), by which one can see that when condition (7) holds, the empty system solution is not stable, while the other steady-state solution is an attractor for all trajectories that do not start from an empty system. In this case, for all initial conditions in which the system is not empty, the system evolves to the same steady-state solution, given by Equation (6). ∎

The significance of the mean availability in steady state a derives from the fact that, in the mean field limit, it coincides with the mean availability of those nodes that enter the ZOI, and hence, with the success probability.

IV. The storage capacity of floating content

As already stated, the floating content paradigm can be seen as a way to implement, through opportunistic replications, a distributed information storage service, enabling *probabilistic* content persistence and retrieval in a geographically limited location, typically without direct support from infrastructure (except possibly for support in initial seeding of the content). In this section we characterize the *storage capacity* of a floating storage system, i.e., the maximum expected amount of information which can be stored probabilistically.

In order to derive a model for such capacity in a way which is analogous to classical information storage systems, we start by considering a single floating element belonging to a floating system (be it localized or distributed). For this system, we define the read and write operations as follows. The write operation consists in the initial seeding of the content within its RZ, with the goal of enabling content to persist probabilistically even after the transient of content diffusion has passed. From Theorem 4 we know that a seeding strategy which attributes each content to at least one node within the RZ enables the system to converge to a nonempty steady state, though in finite systems more conservative seeding strategies are often necessary in order to decrease the likelihood of content disappearance from the system during the initial transient of content diffusion.

In order to define the read operation, we recall that the main purpose of FC is to ensure that a given (minimum) fraction of moving nodes possesses the content item by the time they enter the ZOI, at which point it is assumed that the content is retrieved by the specific application. For instance, the ZOI might correspond to the location of a movie theater, and the content item to a movie trailer which we assume will start being requested when users enter the theater. Therefore, to each node entering the ZOI corresponds a content request, and hence a read operation, which is considered as failed if the node does not possess the given content. The goal of FC is therefore to populate proactively the host local cache in a distributed, collaborative manner, based on the opportunistic content replications between the users which will enter the ZOI and all the other users in the RZ for that content. In information theory, a storage system is modeled as a communication channel [19]. In this view, information is transmitted over the channel through a write operation, and it is received through a read operation. According to the operational definition given by Shannon [20], the channel capacity is the largest amount of bits *per channel use* at which information can be sent on the considered channel with arbitrarily low error probability. For the specific case in which FC is used as a storage technology, a channel use is the operation of setting a content to float in the floating area, i.e., a write operation for a content of size L. Let us denote as *stored information* of a FE the mean amount of information which can be recovered (i.e., read). Then, similarly to communication channels, if we consider the maximum of this quantity over all channel uses (and therefore content sizes), we have the following definition of storage capacity of a floating element:

Definition 2 (**Storage capacity**)**.** The storage capacity of a floating element with radius R is the maximum of the stored information, over all content sizes $L \leq M$.

In what follows, we consider the floating element and its floating system to be in a stationary state. With $a(R, \gamma, L)$ we denote the mean field limit of the mean availability within its RZ for the stored content, for a RZ radius R.

Theorem 5. *In a floating system in the stationary regime, the mean storage capacity of a floating element is*

$$C_{FE}(R, \gamma) = \max_{L \leq M} La(R, \gamma, L), \tag{8}$$

where $a(R, \gamma, L)$ is given by Theorem 4.

Proof: At the mean field limit, every node in the RZ has the same probability to possess the content, independently from its position. Hence, at the mean field limit, the success probability P_{succ} in the given FE coincides with $a(R, \gamma, L)$. In the communication channel model of storage systems, a FE can be modeled as a *packet erasure channel* [21], with packet size equal to the size of the content, L, and packet erasure probability $1 - P_{succ}$, i.e., equal to the probability for a node not to possess the content when it enters the ZOI. In such a model, every channel use is a write operation, consisting in setting a content of size L bits to float in the floating region, and erasures in the channel derive from the fact that the read operation is not deterministic. The amount of bits which can be recovered, on average, on a single channel use is hence $La(R, \gamma L)$. The maximum of this quantity over all channel uses, and hence over all content sizes $L \leq M$, is the storage capacity of the system. ■

One of the consequences of Theorem 5 is that an upper bound to the mean storage capacity of a FE is given by the minimum between the host memory M, and the largest content size which satisfies Equation (7). Indeed, by increasing content size, the amount of information stored increases, but so does the average amount of time taken by content transfers, decreasing the share of successful content transfers and hence the availability for that content. For small values of content size, the first effect prevails. For larger values however, at some point the second effect takes over, decreasing the amount of stored information, until the condition in Equation (7) is not satisfied, bringing the amount of stored information to zero.

In a floating system composed by more than one floating element, the overall amount of information stored is also a function of how the RZs overlap, hence of both FE intensity γ and RZ radius R. The following result gives the capacity per unit area of a floating system for a given FE intensity, and a given RZ area.

Corollary 1. In a floating system (localized or distributed) in the stationary regime, the mean storage capacity per unit area is given by

$$C_{FS}(R, \gamma) = \gamma \max_{L \leq M} La(R, \gamma, L), \tag{9}$$

where $a(R, \gamma, L)$ is given by Theorem 4.

Proof: The system can be modeled as a set of γ parallel, independent packet erasure channels per unit area, one per distinct content. The independence between the channels holds at the mean field limit and it derives from the "propagation of chaos" result, by which at the mean field limit for each user the probability to have a content is independent from the probability of having another content. For each content, the expression of the capacity of the associated packet erasure channel is given by Theorem 5. ■

In a floating system, the RZ radius modulates the average amount of users, hence of system resources, dedicated to a given FE, while the FE intensity tells how many floating contents on average are sustained by the system per unit area. In what follows, we are interested in how to modulate these parameters in order to maximize the average amount of information per unit area stored in a floating system. We have therefore the following optimization problem:

Problem 1: $\underset{R,\gamma,L}{\text{maximize}} \quad \gamma L a(R,\gamma,L)$

$$\text{subject to: } Equation\ (3),\ (4),\ (6),\ (7),$$
$$0 \leq R, \tag{10}$$
$$1 \leq L \leq M, \tag{11}$$
$$0 \leq \gamma \leq \frac{DM}{L}. \tag{12}$$

Constraint (12) derives from the fact that an upper bound to the intensity γ, and hence to the average number of different contents floating in a given area is given by the average number of nodes present in that area, multiplied by the number of contents which each node can store. As a consequence, an upper bound to the amount of stored information per unit area is DM, and it corresponds to the case in which, for each content, a single copy exists in the system, so that the system has no redundancy. As for the content size L, we have:

Proposition 1. If (γ^*, R^*, L^*) is a solution of Problem 1, then $L^* = 1$.

This result derives from the fact that, holding constant all else, the lower the content size, the lower the amount of contact time wasted in content transfers which do not complete due to the end of contact time. Moreover, the lower the content size, the higher the amount of information which each node can store in their finite memory, as the memory size is not always an exact multiple of content size. As a consequence, Problem 1 becomes a maximization problem over R and γ only, with content size equal to its minimum value $L = 1$. Finally, it can be easily shown that, for any choice of the other system parameters, there exist always a RZ radius beyond which availability decreases monotonically with increasing R. Summing up, despite Problem 1 is nonconvex and nonlinear, it is function of two variables over finite intervals, and it can hence be solved efficiently by brute force approaches.

V. Numerical assessment

In this section, we evaluate numerically the accuracy of our model by means of simulations, and we characterize the storage capacity of a floating system as a function of the main system parameters and of node mobility.

Unless otherwise specified, we assume nodes move according to the Random Direction Mobility Model, with reflection at the boundary of the simulation area. When two nodes are in contact, we assume the channel rate is constant over time and equal to 10 Mb/s. Nodes have a transmission radius of 30 m. For localized floating systems, the simulation area is a square of side 500 m, the RZ has a default radius of 100 m and it is located at the center of the area.

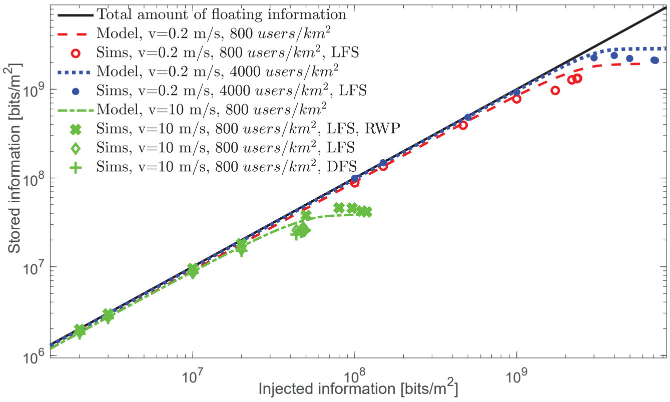

Fig. 1: Information stored per unit area in a distributed floating system (DFS) and localized floating system (LFS) versus injected information per unit area. Simulation are with a 95% confidence interval of at most 0.35%.

At the beginning of each simulation run, nodes are distributed uniformly at random. In order to minimize the probability of content loss during the content diffusion transient, all nodes within the RZ for a given content possess it at set up time. When node memory is finite, the set of contents possessed by each node at set up time is a random subset (different for each node, and of total size equal to the node memory) of all those contents in whose RZ the node is located. Simulated time has been divided into equally sized slots, and their duration has been chosen in such a way as to minimize the effects of quantization in time on accuracy of simulations, and in particular on the errors in detecting when two nodes are in range or when a node is within a given RZ. Content size has been set in such a way to have the start and end of simulation time slots coincide with the start and end of content transfers. Specifically, it has been set to 5 Mb in setups with node speed of 0.2 m/s, and to 100 kb for those with a node speed of 10 m/s. Simulations have been run for a duration of 10000 time slots, which has proven to be sufficient to observe the system out of any transient, and data affected by transient effects (due to initial distribution of nodes, plus possibly due to initial spatial distribution of nodes) have been discarded.

In a first set of simulations, we measured the amount of information stored per unit area in both distributed and localized floating systems, in the ideal case of unlimited host memory. In particular, we characterized how the stored information varies as a function of the *injected information per unit area*, i.e., of the product of the content size and of the mean number of contents per unit area which persist in the system for the whole duration of the simulation. This quantity is modeled by the product γL when condition (7) holds, because in that case all contents float and carry L bits. As Fig. 1 and 2 suggest, two regimes can be observed. For low amounts of injected information and with infinite host memory, resource contention among different floating contents is weak, as the mean contact time is larger than the mean

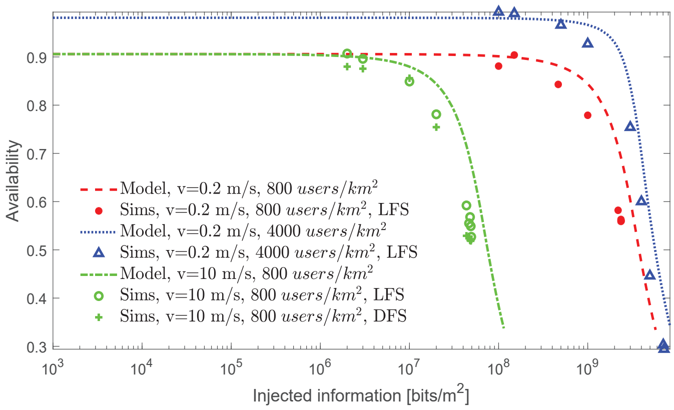

Fig. 2: Mean availability in a distributed floating system (DFS) and localized floating system (LFS) versus injected information per unit area. Simulation are with a 95% confidence interval of at most 0.35%.

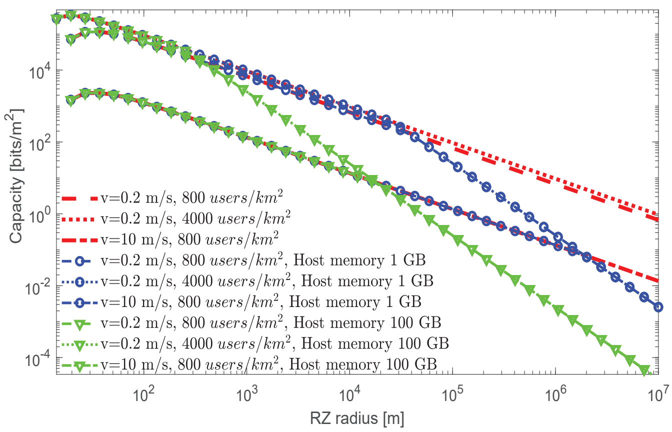

Fig. 3: Maximum storage capacity of a floating system over FE intensity, as a function of RZ radius R, for different configurations of host memory, and with content size $L = 1$.

amount of time required to exchange contents. Therefore, each FE performs almost as if in isolation. Indeed, mean availability remains constant with increasing injected information, while stored information grows proportionally to it. For larger values of injected information, the effects of contention (mainly on contact time) kick in, mean content availability decreases, and stored information saturates. As expected, decreasing the node speed (and therefore increasing mean contact time) and increasing node density (hence also the rate of contacts among nodes, and the opportunities for content replication) have the effect of increasing the saturation value of the stored information, which coincides with the capacity of the system.

The plots in Fig. 1 and 2 also show that the estimates of mean availability and of the amount of stored information derived with our mean field based approach are accurate across different values of node density, of injected information, and of mean node speed. When the amount of injected information gets close to the maximum amount which can be sustained by the system, our mean field model yields slightly optimistic results, due to the difference between finite systems and their mean field limit. Indeed, as the plots show, increasing the mean number of nodes in the RZ improves the accuracy.

A first reason for such discrepancy is that while contact events are uniformly distributed within the RZ, in finite systems nodes with content are slightly more numerous around the center of the RZ than at the border [4]. The overall net effect is a decrease in opportunities for content replications within a RZ. This is also the reason for the slightly more pessimistic simulation results of distributed floating systems with respect to localized ones. Indeed, in systems with distributed RZs, and particularly for low densities of FEs, much of the overlapping involves mainly the border regions of each RZ. As shown in the figures, such difference tends to decrease as the FE intensity grows. In general, the difference between considering a distributed or a localized floating system is small, so that in the following we only report results for LFS.

Another reason for the discrepancy is the effect of random fluctuations in a finite population of nodes. Indeed, when the system gets close to those conditions in which (7) is not satisfied (and hence in which contents cannot persist in their RZs), the effects of even small random perturbations in a finite population of nodes with content get amplified, because of resource contention and the consequent loss of efficiency of the content replication process. This brings to an average decrease in content availability, and to the difficulty in achieving those values of maximum injected information forecasted by the model.

In order to evaluate our approach in non homogeneous scenarios, in Fig. 1 we also report simulation results under the random waypoint (RWP) mobility model, which is known to generate a non uniform distribution of nodes. As expected, mean content availability and hence the amount of stored information are slightly higher than in RD mobility model. Indeed, in RWP nodes from all of the simulation area tend to spend a larger amount of time near the center of the area, where the RZ is. Moreover, if we account in our model the average density of nodes generated in the RZ by RWP, the match between simulation and analytical predictions becomes even better.

One of the key parameters affecting the mean capacity of a floating system is the RZ radius, which is related to the total amount of node memory and content exchanges dedicated to storing probabilistically a single content. Fig. 3 shows the maximum storage capacity of a floating system over FE intensity (Problem 1) as a function of RZ radius. As the figure shows, the solution of Problem 1 is achieved for values of RZ radius only slightly larger than the minimum values below which contents do not persist in the RZ. For lower values of the RZ radius, the content availability, and hence the maximum storage capacity, decrease rapidly, as smaller RZs imply less opportunities for contents to replicate. For values of the RZ radius larger than the optimum instead, the benefits of a larger RZ are offset by the fact that a much larger amount of nodes is involved in content replication, so that the marginal utility of adding more users to each RZ is negative. By further increasing the RZ radius, the maximum stored information reaches a regime where host memory limits start to affect system performance, further decreasing the marginal utility

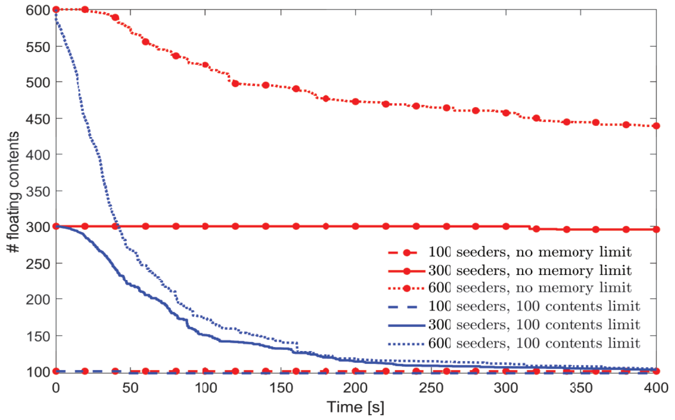

Fig. 4: Mean number of floating contents over time in a localized floating system, for different initial number of seeders. Node speed is equal to 10 m/s, and node density is 800 $users/km^2$.

of adding more users to each RZ.

As we have seen, for a given content size, node density and RZ radius, when the density of contents seeded in a floating system is larger than the maximum amount which can be sustained by the system in steady state, according to our model the system converges to the empty state. However, in finite systems, as Fig. 4 shows, contents start disappearing from the system until the remaining ones can be sustained in the mean field regime according to condition (7). Note that when host memory is finite, if the system is seeded with a higher number of contents than those which can be stored in host memory, the total number of floating contents decreases rapidly during the initial transient of content diffusion, until it coincides with the maximum number of contents which can be stored in host memory.

VI. Conclusion

In this paper, we have developed a first analytical model of the amount of information which can be stored in probabilistic distributed edge storage systems such as Floating Content. Our approach enables a first order characterization of the relation between storage capacity and the main design parameters of such systems, which is crucial for resource efficient and QoS aware dimensioning of such systems.

As a future step, we plan to extend the proposed approach to real settings with heterogeneous shapes and sizes of replication zones, and to scenarios with non stationary mobility patterns and non uniform nodes distribution.

References

[1] T. Xu, J. Jiao, X. Chen, and Y. Chen, "Social-Aware D2D Caching Content Deployment Strategy over Edge Computing Wireless Networks," in *ICCCN*, July 2018, pp. 1–6.

[2] X. Zhang and Q. Zhu, "Collaborative Hierarchical Caching over 5G Edge Computing Mobile Wireless Networks," in *IEEE ICC*, May 2018, pp. 1–6.

[3] S. Zhang, J. Wu, Z. Qian, and S. Lu, "Mobicache: Cellular traffic offloading leveraging cooperative caching in mobile social networks," *Computer Networks*, vol. 83, pp. 184 – 198, 2015.

[4] E. Hyytia, J. Virtamo, P. Lassila, J. Kangasharju, and J. Ott, "When does content float? Characterizing availability of anchored information in opportunistic content sharing," in *IEEE INFOCOM*, April 2011.

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

[6] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *Communications Surveys Tutorials, IEEE*, vol. 16, no. 4, pp. 1801–1819, 2014.

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

[8] B. Liu, B. Khorashadi, D. Ghosal, C.-N. Chuah, and M. H. Zhang, "Assessing the VANET's local information storage capability under different traffic mobility," in *INFOCOM*. IEEE, March 2010.

[9] B. Xie, Y. W. Chen, M. Xu, and Y. G. Wang, "Mathematical modeling of locally information storage capability of VANET for highway traffic," in *Applied Mechanics and Materials*, vol. 373. Trans Tech Publ, 2013, pp. 1914–1919.

[10] G. Manzo, M. Ajmone Marsan, and G. A. Rizzo, "Analytical models of floating content in a vehicular urban environment," *Ad hoc networks*, vol. 88, pp. 65–80, May 2019.

[11] G. Manzo, J. S. Otalora, M. Ajmone Marsan, and G. Rizzo, "A deep learning strategy for vehicular floating content management," *ACM SIGMETRICS performance evaluation review*, vol. 46, no. 3, pp. 159–162, January 2019.

[12] S. Ali, G. Rizzo, V. Mancuso, V. Cozzolino, and M. Ajmone Marsan, "Experimenting with floating content in an office setting," *IEEE Communications Magazine*, pp. 49–54, June 2014.

[13] R. Yamamoto, A. Kashima, T. Yamazaki, and Y. Tanaka, "Adaptive contents dissemination method for floating contents," in *IEEE VTC2019-Fall*, September 2019.

[14] A. Kolesnichenko, V. Senni, A. Pourranjabar, and A. Remke, "Applying mean-field approximation to continuous time Markov chains," in *ROCKS 12*. Springer, 2012, pp. 242–280.

[15] J. H. Argyris, G. Faust, M. Haase, and R. Friedrich, *An Exploration of Dynamical Systems and Chaos - 2^{nd} Ed.* Springer, 2015.

[16] A. Chaintreau, J.-Y. Le Boudec, and N. Ristanovic, "The age of gossip: Spatial mean field regime," in *SIGMETRICS Performance Evaluation Review*, vol. 37, no. 1. ACM, 2009, pp. 109–120.

[17] J. Virtamo, "38.3143 Queuing Theory - Markov processes," https://www.netlab.tkk.fi/opetus/s383143/kalvot/E_markov.pdf, Tech. Rep.

[18] J.-Y. Le Boudec, "The stationary behaviour of fluid limits of reversible processes is concentrated on stationary points," *arXiv preprint arXiv:1009.5021*, 2010.

[19] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[20] C. E. Shannon, W. Weaver, and A. W. Burks, "The mathematical theory of communication (review)," *Philosophical Review*, vol. 60, no. 3, pp. 398–400, 1951.

[21] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 819–828, 2000.

Appendix

A. Proof of Lemma 2

Let us denote with a and b the two nodes in contact. Let m_a (m_b) denote the amount of contents possessed by node a (resp. b) at time t for RZ radius R, and let c_a be the amount of exchangeable contents at node a. The probability that node a possesses exactly c_a exchangeable contents is equal to the probability that c_a out of the m_a contents are not possessed by the other node, and that the remaining contents are possessed by the other node. Moreover, c_a is upper bounded by the available storage space at node b, equal to $\frac{M}{L} - m_b$. c_a is therefore distributed as a binomial with parameters $n = m_a$, $p = 1 - a(t, R)$ and truncated at $c_{a,\max} = \min(m_a, \frac{M}{L} - m_b)$. Specifically, for $0 \leq c_a \leq c_{a,\max}$ we have

$$P(c_a|m_a, m_b) = \frac{\binom{m_a}{c_a}(1 - a(t, R))^{c_a} a(t, R)^{m_a - c_a}}{F_{c_a}(c_{a,\max})} \quad (13)$$

Otherwise, $P(c_a|m_a, m_b) = 0$. The normalization factor $F_{c_a}(c_{a,\max})$ is the CDF of the binomial with parameters $n = m_a$, $p = (1 - a(t, R))$. As m_a and m_b are random variables (whose distribution is given by Lemma 1), the final expression of the PDF of c_a is derived as the expectation of the expression in Equation (13) with respect to m_a and m_b. The PDF of c_b is computed in the same way, and expressions can hence be derived from those of c_a, by substituting m_a with m_b and vice versa. As m_a and m_b have the same distribution, the PDF of c_b is the same as c_a. The expectation of the total number of exchangeable contents on a contact at time t is given by the sum of the expectations of c_a and c_b, i.e. by twice the expectation of c_a.

B. Proof of Theorem 1

First, we show that, for any initial conditions $\mathbf{I}(0, R) = (a(0, R), b(0, R))$, there exists an array $\mathbf{I}_0$ such that $\lim_{R\to\infty} \mathbf{I}(0, R) = \mathbf{I}_0$ (*convergence of initial conditions* condition [14]). Let us choose $\mathbf{I}_0 = \mathbf{I}(0, R)$. Then for each content j, if $N_j(0, R)$ ($N_j^T(0, R)$) is the number of nodes with the j-th content in the RZ of content j (respectively, the total number of nodes in the RZ of content j) at time $t = 0$ in the RZ, choosing $N_j(0, R) = \lfloor a(0, R)N_j^T(0, R)\rfloor$, and setting to zero the number of busy nodes at $t = 0$ allows satisfying the convergence condition. Given the assumption of stationarity of the mobility patterns, and of uniform node distribution, the mean total number of nodes in a RZ for a content j is equal for each content (given that all RZ have the same shape and size) and we denote it with $N(R)$. In order to apply the mean field approximation approach, we start by assuming $N_j^T(t, R) = N(R)$, for any content j an any time $t \geq 0$. As a consequence of the homogeneous condition, $N(R)$ grows proportionally to R. With these properties, the considered system can be modeled as a Population Continuous Time Markov Chain (PCTMC) [14]. Specifically, to each value of R we can associate a PCTMC model with a total number of nodes $N(R)$. As for the *size* of the model (i.e., as for the parameter used for normalizing the state occupancy), we choose the parameter $N(R)$ itself. Let us consider now a sequence of increasing values of R, to which we can associate a sequence of PCTMC models, each with the features described so far. By the nature of the system, one can easily verify that for any state transition, the state change vector (i.e., the difference between the state occupancy before and after the state transition) is independent of R and hence of the size of the model. Let $N_c(t, R)$ denote the number of nodes possessing a given content j at time t, averaged across all j. Let us compute the rate at which $N_c(t, R)$ varies over time. The increase of this quantity is due solely to nodes which exit from the busy state due to completion of content transfers. The mean rate at which nodes exit the busy state is given by the ratio between the mean number of busy nodes at time t in the RZ, denoted with $N_b(t, R) = N(R)b(t, R)$, and the mean time taken by an exchange, $T_s(t, R)$. However only half of the $N(R)b(t, R)$ busy nodes which terminate an exchange might have received the j-th content. Moreover, let us consider one of these terminating exchanges. The probability that the j-th content was transferred during such exchange is equal to the probability that only one of the two exchanging nodes had the j-th content, given by $2a(t, R)(1 - a(t, R))$, multiplied by the probability that the j-th content was transferred during the contact time, given by $S(t, R)$. Putting it together, we have that the rate of increase of $N_c(t, R)$ is $\frac{N(R)b(t,R)}{T_s(t,R)} a(t, R)(1-a(t, R))S(t, R)$. The decrease of $N_c(t, R)$ over time is due to nodes with content exiting the RZ. The overall rate at which nodes exit the RZ is $\alpha 2\pi R$. Of these nodes, only a fraction $a(t, R)$ possesses the i-th content. Summing up, we can write

$$\frac{dN_c(t,R)}{dt} = \frac{N(R)b(t,R)}{T_s(t,R)} a(t, R)(1 - a(t, R))S(t, R) - 2\alpha\pi Ra(t, R) \tag{14}$$

Let us consider now the rate at which $N_b(t, R)$ varies over time. The rate of increase of this quantity is given by the rate of those contacts involving nodes which are not busy. The probability that a contact happens between two non-busy nodes is $(1 - b(t, R))^2$. As every such contact generates two new busy nodes, and as $g\pi R^2$ is the overall contact rate, the rate of increase of $N_b(t, R)$ over time is given by $2g\pi R^2(1 - b(t, R))^2$. The number of busy nodes in the RZ decreases due to busy nodes exiting the RZ, and to the end of the exchange of contents between two nodes. The rate of the first type of events is $4\alpha\pi Rb(t, R)$. Note the extra factor 2, due to the fact that if a busy nodes exits the RZ, both it and the other node in the exchange stop exchanging contents, hence both are no more busy, even if the other node remains in the RZ. Finally, the rate at which nodes cease being busy due to the end of the exchange of contents is given by $\frac{N_b(t,R)}{2T_s(t,R)}$, i.e., by the rate at which couples of exchanging nodes "break" due to the end of the exchanges, multiplied by a factor 2, as at each of these events the number of busy nodes decreases by 2 units. Summing up, we have

$$\frac{dN_b(t,R)}{dt} = 2g\pi R^2(1 - b(t, R))^2 - 4\alpha\pi Rb(t, R) - \frac{N_b(t,R)}{T_s(t,R)} \tag{15}$$

From the rates in these equations it is easy to see that the drift of the generic PCTMC is continuous. Let $\mathbf{I}(t, R) = (a(t, R), b(t, R))$ and $\mathbf{I}(t) = (a(t), b(t)) = \lim_{R\to\infty} \mathbf{I}(t, R)$. As our sequence of PCTMC models satisfies these properties, by Theorem 1 in [14] we have that for any finite time horizon $T \leq \infty$, $\mathbb{P}\left\{\lim_{R\to\infty}\left(\sup_{0\leq t\leq T} \|\mathbf{I}(t, R) - \mathbf{I}(t)\|\right) = 0\right\} = 1$. That is, the sequence of population models associated to R converges *almost surely* to the dynamics of the ODEs in Theorem 1. Finally, in the case in which the number of nodes in each RZ is not constant, one can follow the same approach and derive an additional differential equation for the mean number of nodes in each RZ. Indeed, as we assumed the mobility is stationary, such differential equation would be a balance equation, giving a mean number of nodes in each RZ which does not vary over time, and which is not affected by the evolution over time of the other two variables of the system. Given such decoupling, the mean field approach can be applied separately to the mean number of nodes in each RZ, and to the two variables we have considered so far, obtaining again the ODEs in (5).