A Simple Approximate Analysis of Floating Content for Context-Aware Applications

Shahzad Ali
Institute IMDEA Networks, Madrid, Spain
Universidad Carlos III de Madrid, Spain
shahzad.ali@imdea.org

Gianluca Rizzo *
HES-SO Valais, Sierre, Switzerland
gianluca.rizzo@hevs.ch

Balaji Rengarajan
Institute IMDEA Networks, Madrid, Spain
balaji.rengarajan@imdea.org

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

ABSTRACT

Context-awareness is a peculiar characteristic of an expanding set of applications that make use of a combination of restricted spatio-temporal locality and mobile communications, to deliver a variety of services. Opportunistic communications satisfy well the communication requirements of these applications, because they naturally incorporate context. Recently, an opportunistic communication paradigm called "Floating Content" (FC) was proposed, to support infrastructure-less, distributed content sharing. But how good is floating content in supporting context-aware applications? In this work, we present a simple approximate analytical model for the performance analysis of context-aware applications that use floating content. We estimate the "success probability" for a representative category of context-aware applications, and show how the system can be configured to achieve the application's target QoS. We validate our model using extensive simulations under different settings and mobility patterns, showing that our model-based predictions are highly accurate under a wide range of conditions.

Categories and Subject Descriptors

J.4 [**Human-centered computing**]: Ubiquitous and mobile computing

Keywords

Context-Aware Applications, floating content, anchor zone.

*Gianluca Rizzo was with Institute IMDEA Networks, Madrid, Spain during this work

MobiHoc'13, July 29–August 1, 2013, Bangalore, India.

1 Introduction

The growth of mobile computing, and the pervasiveness of smart user devices is progressively driving applications towards context-awareness, i.e., towards the exploitation of "any information that can be used to characterize the situation of an entity" [7], in order to implement new services that better suit the needs of users. One example of context information is given by spatio-temporal coordinates. For a parking finding application, the information about a vacant parking spot is characterized by the position of the spot, and it is valid for a limited time (until the space is filled). Many more examples of context-aware applications are emerging, that make use of spatio-temporal context and wireless communications to deliver a variety of services. For most context-aware applications, the scope of generated content is limited in time and space. The local relevance of the content generated and exchanged by context-aware applications makes their communication requirements significantly different from those of ordinary applications. Indeed, this locality feature makes traditional centralized, infrastructure-based content distribution services unfit, as they imply a heavily suboptimal use of storage and connectivity resources when supporting context-aware applications. Opportunistic communications can play a special role when coupled with context-awareness. Indeed, they naturally incorporate context and locality as spatial proximity is closely associated with connectivity. In our work we consider a specific opportunistic communication paradigm, known as *floating content* (FC) [3], conceived to support server-less distributed content sharing. It aims at ensuring the availability of some content within a certain geographic area called *anchor zone* (AZ), and for a given time period. Initially, only one node generates the content and defines the AZ for that content. Within the AZ, any time a user who is unaware of the content enters the transmission range of another user possessing it, the content is shared. By being replicated and made available to nodes within the AZ, even after that the original node which generated it leaves the AZ, the content 'floats' within the AZ. Users who traverse the AZ while content is floating have an opportunity to learn the content, provided they meet a node with content prior to leaving the AZ. Users delete content if they move outside AZ. Due to stochastic fluctuations of the number of nodes with content within the AZ, and to their traversal patterns, the content ultimately disappears from the AZ. In the previous modeling studies of FC [3, 5] the focus was on understanding the asymptotic properties of the floating lifetime, i.e. the duration of time for which content floats in the AZ, and asymptotic conditions were derived (called criticality conditions) for the expected floating lifetime to be large under some large population assumption. Instead, our objective is to characterize the performance of context-aware applications relying on the FC service. Indeed, from an application perspective, it is not sufficient that the content asymptotically floats: what matters is the probability for a node traversing the AZ to get the content which is floating. In this paper we address this issue, investigating the effect of the system design parameters on the performance of applications using FC. We restrict ourselves to a regime where floating lifetime is expected to be large, and we focus on the *success probability*, which captures the likelihood for a user to receive the content when traversing the AZ. A key characteristic of our modeling approach is that success probability is computed from few primitive system parameters, most notably the probability density function of the length of the path followed by users within the AZ. This allows our analysis to be generalized to a variety of AZ shapes, user mobility patterns, user speed distributions, and applications. Our main contributions are: i) we develop an approximate analytical model for success probability, with key parameters the AZ radius, the node density, and the node transmission range; ii) we apply our model to a representative category of context-aware applications, and derive expressions for their success probability; iii) we demonstrate how the model predictions can be used to tune key system parameters to achieve the desired application performance; and iv) we validate our analysis with extensive simulations, showing that the predicted success probability is very accurate. The rest of the paper is organized as follows. In Section 2, we present the system model and we define the performance metrics. We present our analysis in Section 3, and in Section 4 we validate it via simulations. Finally, we present our conclusions in Section 5.

2 System Model

We consider mobile nodes in $\mathbb{R}^2$ distributed at any time t according to a homogeneous Poisson point process with intensity λ. This assumption is justified by the considered mobility model.

Mobility Model: We assume that nodes move according to the Random Direction (RD) mobility model [2], in which nodes independently travel along a straight line, with an angle of movement uniformly distributed between 0 and 2π, and at a constant velocity v. In [4] it is shown that under this mobility model, the spatial node distribution remains uniform at all time instants. The reason for choosing this mobility model is its simplicity and analytical tractability.

Content Replication: We assume that at time $t = 0$ we have a "seeder" node with a given information item, which has to be made available to (possibly all) nodes within its *anchor zone* (AZ). AZ is a circular area with radius R, centered at the position of the seeder at $t = 0$. Then for $t > 0$ every time two nodes *within the AZ* come in transmission range of each other (we call this a contact), if only one of the two nodes possesses the item, it transfers it to the other. We assume there is no supporting infrastructure available, so that nodes must rely exclusively on ad-hoc communication. All nodes are assumed to have the same transmission range r, with $R >> r$, since these are the cases where content floating has practical utility. When nodes move outside of the AZ, they delete their own copy of the item. We assume every node knows his exact position in space at any time.

Performance Metric: In this paper, we focus on scenarios where node density and anchor zone radius are such that the criticality condition derived in [3] holds, so that the expected lifetime of content floating is infinite under the fluid limit approximation of [3]. In practice, this implies large average floating lifetimes. The performance metric that we use is the *success probability*:

DEFINITION 1 (SUCCESS PROBABILITY). *Consider an information item floating in its anchor zone. The success probability $P_s(\tau)$ of the item is the probability that a node receives it within a time τ after entering the AZ (if the node is still in the AZ), or by the time it leaves the AZ (if it leaves it before time τ), averaged over the duration of the floating lifetime.*

3 Analysis

In this section we derive an expression for approximating the success probability for the general floating content model described in Section 2. Then, we extend our analysis to a family of applications, and derive approximate expressions for success probability for these applications.

General Floating Content: We assume in our derivation that the system is in an equilibrium state (as assumed in [3]), in which the content floats indefinitely, and in which the average rate at which nodes get the information item inside its AZ equals the average rate at which nodes having the item leave the AZ. In such an equilibrium, the average number of nodes with and without the item within the AZ remains constant.

RESULT 1. *Consider an AZ with radius R, node density λ, and nodes with transmission range r and speed v. Let Q denote the probability that two nodes successfully transfer the content while they are in contact. Then $P_s(\tau)$ for $\tau \le 2R/v$ can be approximated as*

$$P_s(\tau) = \int_0^{2R} \frac{l^2}{\pi R^2 \sqrt{4R^2 - l^2}} \cdot$$

$$\cdot \sum_{k=1}^{\infty} \left[1 - \left(1 - \frac{Q\overline{n}}{(\overline{m} + \overline{n})} \right)^k \right] \frac{(2r\lambda(l \wedge v\tau))^k e^{-2r\lambda(l \wedge v\tau)}}{k!} dl \quad (1)$$

where $\overline{m} = \min(\frac{v}{Q\nu R}, \lambda\pi R^2)$, $\overline{n} = \lambda\pi R^2 - \overline{m}$, and $\nu = \frac{2rv^2}{(\pi R^2)}$.

Here, $a \wedge b$ stands for $\min(a, b)$. $\overline{n}$ and $\overline{m}$ are respectively the average number of nodes with and without content within the anchor zone. For the derivation of Result 1, please see Appendix.

Note that in the expression above, the integral is over l which is the length of the AZ chord traversed by a node. The expression calculates the probability that a node meets k other nodes during its traversal as the product of the pdf of the chord length and the conditional pdf of the number of contacts given the chord length. The term in square brackets is the probability that at least one out of the k nodes met has the content, and that the transfer is successful. In deriving the above result, we assume that the distribution of nodes with content in the AZ is uniform, and that the odds of meeting nodes possessing the information are uncorrelated. In reality, both assumptions are not satisfied, since there is some spatial clustering of nodes with content, and also a higher density of such nodes near the center of the AZ. However, as we show through simulations in the sequel, these are second-order effects and the result in (1) captures the success probability well. In ideal conditions, when node density is sufficiently high and Q is close to 1, we see that $P_s(\tau) \approx Q$, as in those conditions the main limit to the performance of the FC service is due to the effectiveness of the information transfer process.

For some applications of floating content, an important performance parameter is the probability P_s that a node obtains the content before leaving the AZ. An expression for it can be derived from Result 1, by letting $\tau = 2R/v$. In Result 1, all factors influencing the probability of successful content transfer between two nodes in range of each other are captured by the parameter Q. This allows detailed models for information transfer at physical and/or MAC layer to be easily incorporated in the analysis, without changing the structure of the formula. In the following theorem we provide a simple expression for the probability of successful content transfer, which accounts for finite bandwidth availability and transmission errors, assuming that nodes continually retry on failure as long as they stay within transmission range.

THEOREM 1. *If X' is the minimum time necessary for the transfer of the content, the probability of successful transfer is given by*

$$Q(X') = \sum_{k=1}^{\infty} \int_{kX'}^{(k+1)X'} [1 - (1 - S)^k] f_\tau(t) dt \qquad (2)$$

where $f_\tau(t)$ is the pdf of the contact duration under the RD mobility model, given by

$$f_\tau(t) = \int_0^{min(2v, \frac{2r}{t})} \frac{2\omega^3 t^2}{\pi^2 r^2 \sqrt{4r^2 - \omega^2 t^2} \sqrt{4v^2 - \omega^2}} d\omega \qquad (3)$$

and S is the probability of no transmission failures (errors, collisions, etc) for each content transfer attempt.

For the proof of Theorem 1, please refer to [1].

Applications performance: The anchor zone defined so far plays a double function. On one side, it is the area within which the content should be available to users, and it is used as a reference for the computation of the success probability. On the other, it is the area within which the content replication mechanism, which makes the content float, is active. To accomplish the first function, in some applications the AZ could be so small to make it impossible for the content to float. In order to better tune system (rather, application) performance, we separate these two functions, and we define the Range Of Interest (ROI) as the zone (circular, and centered at the same point as the AZ) within which the content should be available to users. Necessarily, the ROI is contained in the AZ, as otherwise the content would not be available in some portions of the ROI. Here we focus on a family of applications for which the content must be delivered to users by the time they leave an area, because the content is expected to trigger some specific actions once they leave that area. One example of such application can be targeted advertising, when users visiting a given site should be notified about some offer/discount (e.g. people in vicinity of a shopping mall notified about some ongoing special offers). The expression for success probability for this kind of applications is derived in Result 2 using the following definition:

DEFINITION 2. *The success probability for getting content before leaving the ROI (P_{SBL}) is the probability that a node entering the ROI gets the content before exiting the ROI, averaged over the duration of the floating lifetime.*

RESULT 2. *For an AZ with radius R_2 and a ROI with radius $R_1 \leq R_2$, P_{SBL} can be approximated as*

$$P_{SBL} = \sum_{k=1}^{\infty} \left(\int_{\sqrt{R_2^2 - R_1^2}}^{R_1 + R_2} \left(f_{L_1}(\ell) \frac{(2r\ell\lambda)^k e^{-2r\ell\lambda}}{k!} \right) d\ell \right) \left[1 - \left(1 - \frac{Q\overline{n}}{(\overline{m} + \overline{n})} \right)^k \right] \qquad (4)$$

with $f_{L_1}(l) = \frac{2\sqrt{R_2^2 - (g^{-1}(l))^2}\sqrt{R_2^2 - (g^{-1}(l))}}{(\int_0^{R_1} g(y)dy) g^{-1}(l)}$ *and* $g(y) = \sqrt{R_2^2 - y^2} + \sqrt{R_1^2 - y^2}$.

For the derivation of Result 2, please refer to [1].

4 Simulations and Results

In this section, we present and discuss numerical results obtained from our analysis, and from simulations. The goal here is, on one hand, to validate the approximate expressions derived for success probability, showing that they are accurate under varied conditions. On the other hand, to show the effectiveness of our analysis in selecting the floating content parameters for a family of applications and under different scenarios. For all simulation experiments, OMNeT++ based framework called INET [6] is used. Confidence intervals at 95% confidence level were evaluated for all cases. The transmission range is $50m$ and nodes move with a constant speed of $10m/s$. We measure success probability in each instance by averaging over the floating lifetime or until a maximum of $50000s$ have elapsed. Fig. 1 shows both the analytical predictions and the empirically determined values of success probability as a function of the AZ radius. In addition to the random direction mobility model (RDMM), we evaluate success probability under the Generalized Manhattan Mobility Model (MGMM) with block sizes of 100m × 150m, and probability of turning left or right at an intersection set to be 0.25 each. To model success probability for MGMM, we empirically measure the average number of nodes met by a node while traversing the AZ as well as the overall rate of contacts. We assume, as we did in the analysis of the RDMM, that node contacts are uniformly distributed in space. In order to compute the predicted success probability, we use an analog of Result 1 with the number of contacts during a node traversal modeled by a Poisson random variable parametrized with the empirical mean.

It can be seen that the model predictions match very well with the simulated results for both RDMM and MGMM, suggesting that the model indeed captures successfully the first order-effects on success probability. The curves also show, as expected, that an increase in both node density and AZ radius improves the success probability, as they both increase the chances of meeting a node having content. For

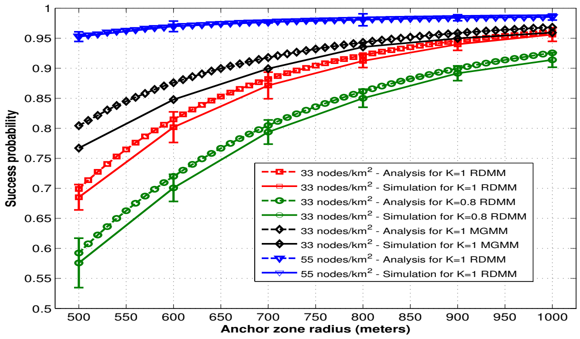

Figure 1: Success probability vs. AZ radius.

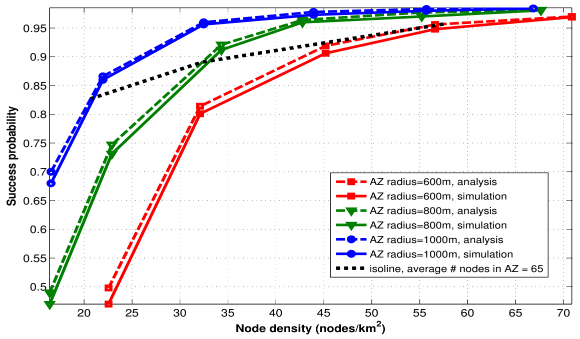

Figure 2: Success probability vs. node density.

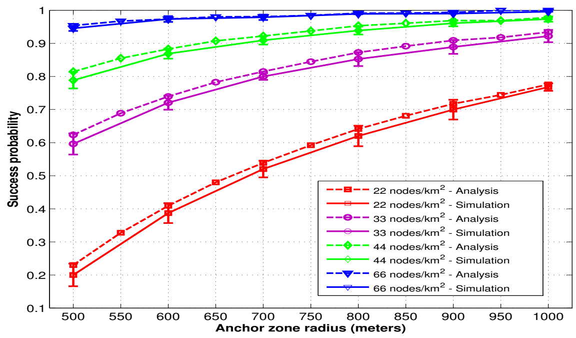

Figure 3: Application success probability with $ROI = 200m$.

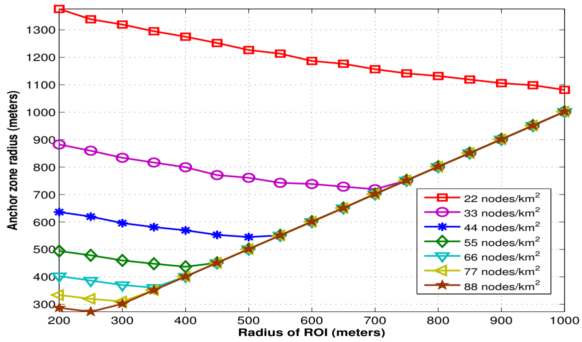

Figure 4: AZ radius to achieve 90% application success probability.

identical node density (33 nodes per km^2), under MGMM the success probability is higher than RDMM. The major reason behind this is the higher contact rate under MGMM compared to RDMM. This increases the population of nodes with content in the AZ, and also the odds of a node meeting a node with content under MGMM. Fig. 1 also shows the impact of transmission errors and finite bandwidth for RDMM. It can be seen that under a finite bandwidth model with a data rate of 11 Mbps, for transmission error probability of 0.2 with a file size of 2MB, the success probability decreases. As the transmission errors and limited contact times reduce the rate of contacts where communication is successful, this reduces the fraction of nodes in the AZ with content as well as the overall success probability. Fig. 2 shows success probability versus node density for different choices of AZ radius. As node density increases, a large increase in success probability can be observed. The analytical model captures this effect, and is a very good predictor for success probability. Further, Fig. 2 shows that as AZs grow larger, a given success probability threshold can be achieved at lower node densities. This is due to node paths through the AZ getting larger with AZ size, resulting in more opportunities to obtain content. However, if node densities and AZ size are varied jointly such that the average number of nodes in the AZ stays unchanged (see the isoline corresponding to an average of 65 nodes in the AZ), larger AZs result in lower success probabilities, since nodes have less chances to come in range of each other and exchange content. Fig. 3 shows curves of success probability versus the AZ radius, for different node densities, for the considered family of applications, when ROI radius is $200m$. Even in this case, increases in either AZ radius or node density result in increased success probability for applications. An interesting feature shown by the curves is that increasing AZ radius has diminishing returns. Indeed, increasing the radius of the AZ affects the density of users with content in the AZ, which is known to be quite flat away from borders [3]. The figure also shows that our analytical results are very close to simulations results. The proposed analytical model can be used to tune AZ radius to achieve a target application success probability. In Fig. 4 we assumed the threshold to be 90%, and used our models to compute the minimum required AZ radius at different ROIs and node densities. Interestingly, for a given node density, as the radius of ROI increases, the required AZ radius decreases, until the condition is reached where replicating the content within the ROI is sufficient to achieve the desired success probability.

5 Conclusions

In this paper, we focus on floating content as a communication service supporting context-aware applications. We defined success probability as the primary performance indicator, and developed a simple, approximate analytical modeling framework, which can be adapted to several different settings. Our models can be adapted to several different categories of context-aware applications, and the model predictions can be used in order to tune key parameters of the system to achieve the required performance with minimum overhead. We show by simulation that our approximation computes very accurate success probability values for a wide range of anchor zone radii and node densities demonstrating the viability of floating content as an enabler for context-aware applications.

References

[1] S. Ali, G. Rizzo, B. Rengarajan, and M. Ajmone Marsan. A simple approximate analysis of floating content for context-aware applications. Technical report, 2013, Available at: http://eprints.networks.imdea.org/511/.

[2] C. Bettstetter. Mobility modeling in wireless networks: categorization, smooth movement, and border effects. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(3):55–66, July 2001.

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

[4] P. Nain, D. Towsley, B. Liu, and Z. Liu. Properties of random direction models. In *INFOCOM '05*, volume 3, pages 1897 – 1907, March 2005.

[5] J. Ott, E. Hyytia, P. Lassila, T. Vaegs, and J. Kangasharju. Floating content: Information sharing in urban areas. In *PerCom 2011*, pages 136 –146, March 2011.

[6] A. Varga. The omnet++ discrete event simulation system. *ESM 2001*, June 2001.

[7] A. Zimmermann, A. Lorenz, and R. Oppermann. An operational definition of context. In *CONTEXT'07*, pages 558–571, 2007.

Appendix

Proof of Result 1: In order to prove Result 1, we need to introduce the following lemmas.

LEMMA 1. *Under the RD mobility model with node density λ, when nodes have a transmission radius of r, and velocity equal to v, the number of contacts made by a node in a time interval τ is Poisson distributed with mean $\mu_C = 2rv\tau\lambda$.*

PROOF. For the proof, please refer to [1]. □

LEMMA 2. *In equilibrium state, for a node traversing an AZ under the Random Direction mobility model, the probability of meeting k nodes is given by*

$$\int_0^{2R} \frac{\ell^2}{\pi R^2 \sqrt{4R^2 - \ell^2}} \frac{(2r\ell\lambda)^k e^{-2r\ell\lambda}}{k!} d\ell \tag{5}$$

PROOF. Due to the properties of the RD mobility model, node distribution is uniform in space at any point in time [4], and as the nodes move in a straight line therefore the pdf of chord length L inside AZ is given by

$$f_L(\ell) = \frac{\ell^2}{R^2 \pi \sqrt{4R^2 - \ell^2}} \tag{6}$$

with $\ell \in [0,\ 2R]$.

Using Lemma 1 with $l = v\tau$, the probability of meeting k nodes along this trajectory can be expressed as a Poisson distribution, with intensity $2r\ell\lambda$:

$$P(meet\ k\ nodes|\ell) = \frac{(2r\ell\lambda)^k e^{-2r\ell\lambda}}{k!} \tag{7}$$

The probability that a node meets k other nodes during its traversal is expressed as the product of the pdf of the chord length and the conditional pdf of the number of contacts given the chord length ℓ. Therefore, the probability of meeting k nodes is given by (5). □

LEMMA 3. *Consider an AZ in equilibrium state, with an average number of nodes equal to $\overline{N}$, and let $\overline{n}$ and $\overline{m}$ denote the average numbers of nodes with and without content respectively. Then $\overline{m} = \min(\frac{v}{Q\nu R}, \lambda\pi R^2)$ and $\overline{n} = \lambda\pi R^2 - \overline{m}$, with ν given by $\frac{2rv^2}{(\pi R^2)}$.*

PROOF. For a detailed proof, please refer to [1]. We denote with $n(t)$ and $m(t)$ the number of nodes with and without content, respectively, at a given time t, and define $N(t) = n(t) + m(t)$. All these quantities vary over time, as nodes move in and out of the anchor zone, and as content is exchanged. We build a set of differential equation which describe how these quantities vary over time. $n(t)$ varies over time: due to (1) nodes with content exiting the AZ and (2) nodes without content in the AZ getting the content. After computing both of these quantities we get

$$\frac{dn(t)}{dt} = \nu n(t)(N(t) - n(t))Q - \frac{n(t)v}{R} \tag{8}$$

The number of nodes inside the AZ without content varies in time due to: (1) nodes without content exiting the AZ, (2) new nodes entering the AZ, and (3) nodes without content in the AZ getting the content, by meeting other node(s) with content. After computing these quantities we get

$$\frac{dm(t)}{dt} = \lambda v\pi R - \nu n(t)(N(t) - n(t))Q - \frac{m(t)v}{R} \tag{9}$$

Since we assume that the system is in equilibrium, the time averages of $m(t)$ and $n(t)$, indicated respectively as $\overline{n}$ and $\overline{m}$, remain constant over time. We can write for $\overline{n}$ and $\overline{m}$ differential equations very similar to those derived above, and we can set both $\frac{d\overline{m}}{dt}$ and $\frac{d\overline{n}}{dt}$ equal to zero. Solving for $\overline{n}$ and $\overline{m}$ we get the expressions in the lemma. □

LEMMA 4. *With the same assumptions as before, the probability that a node gets content given that it meets k nodes is given by*

$$1 - \left(1 - \frac{Q\overline{n}}{(\overline{m} + \overline{n})}\right)^k \tag{10}$$

PROOF. The probability for a node to successfully get the content upon meeting another node, can be computed by the product of the probability that the encountered node has the content, (equal to the average fraction of nodes having content in the AZ, and given by $\frac{\overline{n}}{\overline{n}+\overline{m}}$), and the probability of successful information transfer Q. The probability in the lemma is then derived as the probability that at least one out of k encounters with other nodes results into a successful transfer of content. □

PROOF. (Result 1) Consider a node traversing the AZ. This node gets content if, during its traversal: (1) at least one of the encountered nodes has the content, and (2) the information is transferred successfully between the two nodes. In steady state, the success probability P_s can be written as

$$P_s = \sum_{k=1}^{\infty} P(meet\ k\ nodes) P(get\ content|\ meet\ k\ nodes) \tag{11}$$

The probability of meeting k nodes is given by Lemma 2, while the probability to successfully get the content upon meeting k nodes is given by Lemma 4. Substituting, we get (1). □