Characterizing Content Sharing Properties for Mobile Users in Open City Squares

Michael Solomon Desta†, Esa Hyytiä†, Jörg Ott† and Jussi Kangasharju‡
†Department of Communications and Networking
Aalto University, Finland
‡Department of Computer Science
University of Helsinki, Finland

***Abstract*—For mobile delay-tolerant networks, different mobility models have been utilized to assess the performance of routing algorithms and applications. Substantial work has gone into understanding the contact characteristics of mobile users to allow evaluation under conditions that approximate the real world. One important finding has been recognizing that contacts humans make at a macroscopic level is derived from daily routines and social interactions so that analyzing social network structures can assist in determining, e.g., suitable peers for message forwarding. While mid- to long-term social interaction patterns assist in delivering messages at larger scales, such patterns become immaterial when considering a microscopic scale such as content sharing in a city square. At microscopic scales, we face an "open" simulation area in which nodes enter and leave rather than a closed system with a fixed set of nodes. Moreover, small scales show more dynamics (e.g., in terms of node density) and steady state analyses become insufficient.**

In this paper, we investigate the operation of a content sharing application, Floating Content, under such microscopic mobility conditions and characterize its behavior for city squares. For its validation, we derive a mobility model for open squares to which mobile nodes enter and from which they depart.

I. Introduction

Opportunistic ad hoc content sharing has seen a substantial amount of research in the past 5–10 years and is becoming an alternative to infrastructure-supported location-based services and content sharing. Opportunistic ad hoc content sharing has several attractive properties that give benefits over infrastructure-supported variants [1] [2]. There is more privacy for users since they do not need to constantly inform a service provider about their location as would be the case when content is hosted on a server in the Internet. A fair amount of content is also very local and ephemeral in nature, lending support to the idea of keeping the content local to where it was published and where it will be needed. Moreover, in infrastructure based systems, issues of *connectivity* (cf. high roaming cost or unavailability of data services) and *temporal validity* may arise. Temporal validity refers to the problem of WORN (Write-Once, Read-Never), where a centrally stored content, often valid for a short period of time is actually left unread and undeleted when it is no longer needed. Many different variants of opportunistic content sharing have been proposed in research over the years, e.g., [1]–[7]. Conceptually they all follow the same communication paradigm, but differ somewhat in lower-level details in terms of implementation and level of detail presented. Content to be published is assigned an area of relevance (anchor zone) and optionally a time-to-live. Users carry mobile devices and when they meet other users, the devices automatically replicate information assigned to the corresponding area. In this paper, we take floating content [1]–[3] as a concrete example and expand on our previous work in new, unexplored directions.

This work was supported by the Academy of Finland in the RESMAN project (grant no. 134363).

Previous work on opportunistic content sharing has mainly focused on evaluating the system performance on a *macroscopic level*, i.e., looking at the overall system performance and its ability to keep the content floating. One key result is the criticality condition [2], which relates the mean number of contacts during a visit in the anchor zone to the probability of content floating. In this paper, our focus is on understanding how floating content would perform at a *microscopic level*, such as an intersection or a city square. At this level, the previous techniques are no longer reliable and we need new ways of estimating the probability of content floating.

Our work is different from past work in two important ways. First we give a detailed analysis of floating content for small systems, including the initialization phase. Second, we study mobility in squares, validate the model with real-world data, and extend earlier results on floating to smaller-scale scenarios.

Specifically, our contributions are the following.

- We refine the criticality condition from [2] to cases with a small number of nodes. The number of contacts required is considerably higher than in large systems [2].
- We study the bootstrapping of floating content and identify necessary conditions to meet expected lifetimes.
- We define and analyze a mobility model for capturing the movement of people in open city squares where people enter the square, move around for a while, and then leave.
- We collect movement data via public webcams from three different squares to validate our analytical results. We show that the more stringent conditions for small systems defined in this paper are typically met and that content would float successfully in these real-life scenarios.

This paper is organized as follows. Section II gives the background. In Sections III and IV, we analyze the initial transient phase in non-spatial fluid and discrete models, while Section V focuses on a spatial model. Section VI presents a new city square mobility model, and Section VII its empirical validation via real-world data. Section VIII discusses the implications of our results, and Section IX concludes the paper.

II. Floating Content

We now present the basic concept and operation of floating content. We refer interested readers to [1]–[3] for a more detailed description of the floating content concept.

A. Basic Concept

Consider a geographical region (anchor zone) where mobile users (nodes) enter, spend some random time, and then leave. A user with a content wants to share the content with visitors of the specific region, but without the use of any infrastructure. Mobile devices with enough storage capacity and wireless interfaces can be used to post and replicate content. This seed node "posts" the content, and interested nodes in the anchor zone get copy of it upon coming in contact with a node possessing it. Thus, content replicates to new nodes epidemically, even when the original node has left the area.

Interested nodes keep copies of items floating around the anchor zone by replicating them whenever they meet. Hence, floating content does not rely on any infrastructure. As all nodes (including the seeder node) eventually leave the anchor zone and delete the content, the availability of content is probabilistic, and the system is a best effort system. When no node in the region has the content, the information has been lost irreversibly (content has "sunk"). Content items tagged with a time to live (TTL) are discarded upon expiry.

B. System Operation

Initially, a node with a piece of content generates an information item I and posts it in the anchor zone. Along with the actual content, item I will communicate anchor zone information (replication and deletion region), content life time, and metadata to make it easier for other nodes to filter and search for content they are interested in. Since the system does not use any existing infrastructure, the seeder must be physically in the anchor zone when posting content. Content starts to replicate in the anchor zone when other nodes come in contact with the seeder node. In general, when two nodes A and B meet in the anchor zone of an item I, and A has I while B does not, then A replicates I to B. Nodes leaving the anchor zone are free to delete their copy of the item. In this paper, we assume that content is always deleted outside the anchor zone, and the replication occurs only inside the zone.

C. Criticality and Finite Systems

Since content replication depends on the transmission range and the locations of the nodes, as well as their mobility, it is not clear in general when floating content can be supported.

As already mentioned, the *criticality condition* defines, for large systems (at fluid limit), the conditions (in terms of node density, transmission range, anchor zone etc.) under which a population of mobile nodes can support the floating content. In a non-spatial black-box model, the condition takes form [2]

$$\frac{2R}{\mu} > 1 \quad \Leftrightarrow \quad \frac{N\nu}{\mu} = m > 1, \tag{1}$$

where R is the total contact rate of nodes in the anchor zone, μ the node arrival/departure rate, N the (mean) number of nodes in the anchor zone, ν the rate at which an arbitrary pair of nodes meets each other in the anchor zone, and m the mean number of contacts per visit. Models taking into account the spatial dimension can be found from [2].

In reality, an anchor zone may be small (i.e., the content is highly localized), be occupied by a small number of nodes, and stochastic fluctuations in the node population have to be taken into account. In fact, a floating content in the real world is constantly in a transient towards extinction, which is bound to happen latest during a night. However, each content item has also a user defined finite lifetime and, in best-effort fashion, it is sufficient that, with a sufficiently high probability, the published content remains available until it expires.

Content survival in a large system relies on averages as stochastic fluctuations become neglible in the macroscopic scale. However, this does not apply to single nodes and at the criticality threshold only a small fraction of the nodes have the item. In particular, when new content is created, it is carried by a single node and it is of uttermost importance that the seeder manages to pass the content item to several nodes. Otherwise, the content is likely to disappear when the seeder leaves the anchor zone. We refer to this phase as *bootstrapping*.

Moreover, a criticality condition tells us only if floating content is sustainable, i.e., if first succesfully distributed across the anchor zone, then the content will *exist* for long periods of time (infinitely at the fluid limit) whenever the criticality condition is met. However, in practice it is rarely sufficient that the content is carried by a small fraction of nodes somewhere in the anchor zone. Instead, we rather require that the content must be "easily" available, which is achieved when a certain strictly positive fraction of the nodes carry it.

Therefore, in this paper we analyze necessary conditions for

1) increasing the proportion of nodes having the content in order to guarantee *availability* and
2) having a reasonably high probability of *bootstrapping* the system, i.e., avoiding an early extinction of the content during the initial phase.

These objectives are not orthogonal, but rather two sides of the same coin. In the following sections, we study different models for floating content, each of which captures the essential dynamic characteristics of such a system. In particular, we show that in order to meet the above two objectives one has to operate clearly above the criticality threshold.

III. Non-spatial Black-Box Model

In this section, we study a non-spatial model for the number of information carrying nodes given in [2]. In this system, the number of nodes in the anchor zone, denoted by N, is assumed to be so large that random fluctuations in N can be neglected. Moreover, the mean sojourn time of nodes is $1/\mu$ and ν denotes the the frequency at which a pair of nodes come in contact with each other. For simplicity, the system is further assumed to constitute a Markov process. Letting p denote the fraction of nodes carrying the content, the net growth rate for

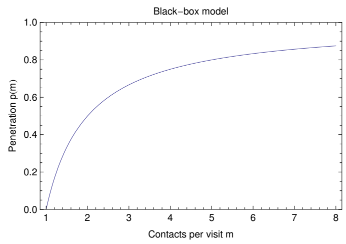

Fig. 1. Mean penetration p according to (4)

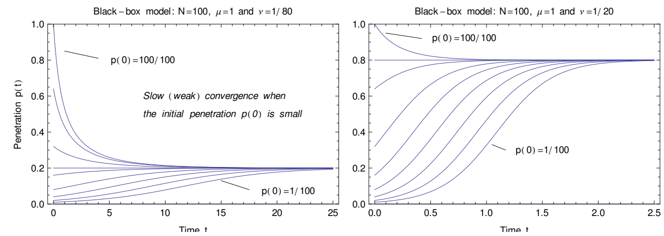

Fig. 2. Solutions to (2) with different initial values $p(0)$.

$p = p(t)$ satisfies the following differential equation [2]:

$$N \frac{dp}{dt} = N^2 p(1-p)\nu - Np\mu. \tag{2}$$

The first term on the right-hand side corresponds to the rate at which content is replicated and the second term to the rate at which information carrying nodes leave the anchor-zone. When dp/dt is negative, the content tends to sink and when dp/dt is positive the fraction of information carrying nodes tends to increase. At a steady state dp/dt is zero.

A. Information availability

At the steady state $dp/dt = 0$ and (2) gives the necessary condition for reaching a penetration higher than p

$$\frac{N\nu}{\mu} = m > \frac{1}{1-p}, \tag{3}$$

which gives (1) when $p \to 0$ and only a marginal fraction of nodes have the information. Conversely, we have:

$$p < \frac{m-1}{m} \tag{4}$$

Fig. 1 depicts the fraction p at the steady state as a function of m. At the criticality threshold $m \to 1$ (from above), i.e., each node meets only one node on average before leaving the anchor-zone. That is, $p \to 0$ and most nodes visiting the anchor zone cannot obtain the content even though it does exist somewhere there. In order to achieve a reasonable *availability*, as stated in Section II-C, we require that the fraction of nodes having the content is strictly positive, $p > 0$ in (3). For example, penetration of $p>0.75$ is achieved if $m > 4$, which is four times higher than what is needed to sustain the content in a large system (given by (1)).

Similarly, the number of *information carrying* nodes a node meets during a visit obeys Poisson distribution with mean $mp = m-1$ (in steady state $m \geq 1$). The necessary condition for *acquiring information with probability of* z then reads

$$m \geq 1 - \ln(1-z). \tag{5}$$

For example, $z = 0.95$ gives $m > 3.996 \approx 4$, which corresponded to 75% penetration (see above).

B. Transient analysis - bootstrapping the system

Let us next focus on the *initial transient*. The differential equation (2) describes the mean behavior, while the actual system has random fluctuations that could be taken into account by analyzing an appropriate stochastic differential equation. However, for our purposes (2) is sufficient.

Suppose that $N = 100$, the mean sojourn time is normalized to one, $1/\mu = 1$, and the rate of pair-wise contacts, ν, is a free parameter. Fig. 2 depicts solutions for the penetration $p(t)$ as a function of time t with different initial values $p(0)$. In the left figure, ν is chosen in such a way that $\lim_{t\to\infty} p(t) = 0.2$, i.e., 20% of the nodes (on average) has the content. In the right figure, 80% of the nodes acquires the content. The lowest curve corresponds to the case that only a single node initially has the content, $p(0) = 1/N$, and the highest curve to the case with $p(0) = 1$. Note that at the criticality threshold $p(t) \to 0$. With $p = 0.2$, the initial transient from $p(0) = 1/N$ is relatively long, about $25 \cdot 1/\mu$, before the steady state is reached. During this time the danger of extinction due to the random fluctuations is considerable; and closer to criticality threshold one is the higher the risk grows. With $p = 0.8$ the situation is clearly better (note the change in time scale).

We conclude that *in order to avoid extinction during the bootstrapping phase, the steady-state penetration* p *should be clearly above the zero*. In other words, the content availability and bootstrapping criteria go hand-in-hand.

IV. Bootstrapping Analysis

In this section, we investigate the floating content during the initial transient phase, where, due to random fluctuations, there is a high chance of content extinction (failure).

A. Probability of Content Absorption

As discussed earlier, floating content requires a sufficient number of nodes in the anchor-zone. However, upon the content creation, only a single node has it and the system is prone to an early extinction due to the stochastic fluctuations.

We model the bootstrapping phase by a *continuous time Markov birth-death process* with $N{+}1$ states, where N is the mean number of nodes in the anchor zone and in state S_i i nodes have a copy of the content. The absorbing state S_0, where the content has *sunk*, is considered as a *failure*, and state S_N as a *success*. The replication of content corresponds to a birth ($S_i \to S_{i+1}$), and an information carrying node leaving the anchor-zone to a death ($S_i \to S_{i-1}$).

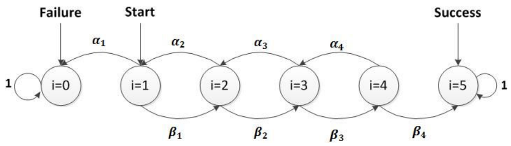

Fig. 3. Floating Content as Birth-Death Process

Let β_i denote the rate at which the number of information carrying nodes grows in state S_i, and α_i the rate at which it decreases. Let U_i denote the probability of absorption into state S_0 when system is initially in state S_i (see Fig. 3). The transition probabilities of the embedded Markov chain are

$$P_{i,i+1} = \frac{\beta_i}{\beta_i + \alpha_i}, \quad \text{and} \quad P_{i,i-1} = \frac{\alpha_i}{\beta_i + \alpha_i}$$

Applying the first step analysis, U_i can be written as:

$$U_i = \frac{\beta_i}{\beta_i + \alpha_i} U_{i+1} + \frac{\alpha_i}{\beta_i + \alpha_i} U_{i-1},$$

where $i \geq 1$ and $U_0 = 1$. As we are interested in the probability of absorption when the system is initialized, U_1, we obtain

$$U_1 = \frac{\sum_{i=1}^{N-1} \rho_i}{1 + \sum_{i=1}^{N-1} \rho_i}, \tag{6}$$

where ρ_i is given by:

$$\rho_i = \frac{\alpha_1 \alpha_2 ... \alpha_{i-1} \alpha_i}{\beta_1 \beta_2 ... \beta_{i-1} \beta_i} = \prod_{k=1}^{i} (\frac{\alpha_k}{\beta_k}). \tag{7}$$

In our case, (2), we know that:

$$\beta_i = N^2 p_i (1 - p_i) \nu, \quad \text{and} \quad \alpha_i = N p_i \mu,$$

where p_i is the proportion of information carrying nodes at state S_i, i.e., $p_i{=}i/N$, giving $\beta_i = (Ni - i^2)\nu$ and $\alpha_i = i\mu$. Hence, $\alpha_i/\beta_i{=}(\mu/\nu)/(N{-}i)$ and substituting this to (7) gives

$$\rho_i = \prod_{k=1}^{i} \frac{\mu}{\nu} \frac{1}{N-k} = \left(\frac{\mu}{\nu}\right)^i \frac{(N-1-i)!}{(N-1)!}. \tag{8}$$

a) Example: Consider a small scenario illustrated in Fig. (3) with only $N{=}5$ nodes in the anchor zone. Substituting (8) with $N = 5$ to (6) gives

$$U_1 = 1 - \frac{24}{24 + z(6 + z(2 + z + z^2))},$$

where $z = \mu/\nu$. At the criticality threshold, the derivative of (2) is zero, $\mu{=}4\nu$, and the above gives

$$U_1 = 47/50 = 0.94,$$

i.e., a system operating at the criticality threshold has a very high failure probability.

Suppose next that $\mu{=}\nu$ so that (2) has a positive derivative indicating a net growth in content density, and we obtain

$$U_1 = 5/17 \approx 0.29,$$

which is a more reasonable value. These simple examples with a *small system size* show that in order to prevent the content from sinking early, a succesful bootstrapping is crucial.

V. Analysis of a Spatial System

As discussed already, in a large system, the responsibility of storing and disseminating the content is effectively shared among a large number of nodes and thus it is sufficient that content replication occurs at a certain *mean rate*. That is, the nodes are *collectively responsible* for the operation. In contrast, when the system comprises a small number of nodes, averages are no longer sufficient but instead each node carrying the content is *individually responsible* for replicating it. Unfortunately, for small systems no obvious and comprehensive single criterion exists, but one has to choose some meaningful objective (see earlier sections). One such first degree objective is to say, e.g., that *the source node must be able to replicate the content at least to two other nodes*. We note that the "thread of life" with an initially empty system (no other node than the seeder has the content) is indeed thin, and the above requirement addresses this issue directly.

A. Direct mobility model

Similarly as in [2], consider a circular anchor zone with radius R. Nodes move in straight line across the anchor zone and all directions are equally likely (isotropic). The node density is denoted by n, so that the mean population in the anchor zone is $\mathrm{E}[N] = n\,\pi R^2$, and the transmission range is d. In this setting, the criticality condition for sustaining floating content at (i.e., a large system with $d \ll R$ reads [2]

$$ndR \geq 0.407. \tag{9}$$

In contrast to the previous models, this model takes into account the spatial dimension explicitly.

B. Bootstrapping analysis

In this model, the contact rate per unit distance is [2]

$$\tilde{\lambda} = \frac{8}{\pi} nd.$$

Assuming the source node starts from the center of the anchor zone and moves directly out (in random direction), the number of nodes it meets before the exit obeys a Poisson distribution,

$$A \sim \text{Poisson}(\frac{8}{\pi} ndR),$$

with mean $\mathrm{E}[A]{=}\frac{8}{\pi}ndR$. At the criticality threshold (9) $ndR{=}0.407$, and we have $\mathrm{E}[A] \approx 1.04$ and $\mathrm{P}\{A > 1\} \approx 0.28$. In particular, $\mathrm{P}\{A = 0\} \approx 0.35$. That is, an initially "empty" system goes to extinction almost immediately with a high probability even at the fluid limit! However, if we require at least two contacts before the exit with a high probability, the resulting system has much higher chances of *bootstrapping* succesfully to the right direction. For example, requiring that

$$\mathrm{P}\{A > 1\} \geq 0.9$$

yields,

$$ndR \geq 1.527,$$

which is over three times higher than (9). This is illustrated in Fig. (4). On the x-axis is the dimensionless quantity ndR and

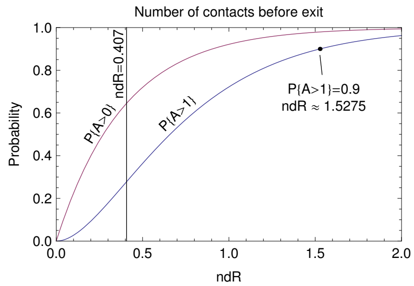

Fig. 4. Requirement that the seeder replicates the content to at least two nodes, $\mathrm{P}\{A>1\}$, as a function of ndR in the direct mobility model.

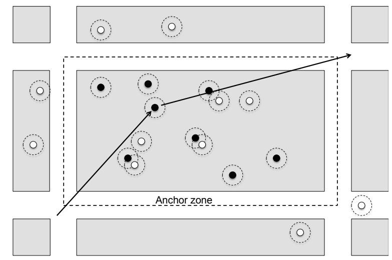

Fig. 5. A rectangular anchor zone with a node entering at the bottom left, changing direction once, and then exiting at the top right corner. "Black" nodes carry a copy of the content (only inside the zone), white ones do not.

the y-axis is the probability. The lower curve corresponds to the probability that the seeder meets at least two nodes before exiting, $\mathrm{P}\{A>1\}$ and the higher curve to the probability that the seeder meets at least one node, $\mathrm{P}\{A>0\}$. We observe that at the criticality threshold of $ndR=0.407$, both events $(A>0)$ and $(A>1)$ are possible but not "certain". However, e.g., at about $ndR \approx 1.5$ the situation is clearly such that bootstrapping the floating is highly probable. The mean length of a visit is $\ell=(\pi/2)R$ [2], and thus the mean number of contacts per visit is $m=4ndR$. With $ndR \approx 1.5$, we thus have $m \approx 6$ nodes and $\mathrm{E}[A] \approx 4$.

Thus, our conclusion is similar as before. *In order to bootstrap the floating with a reasonably high probability, the mean number of contacts, by the seeder and a random node, should be higher than what the criticality condition suggests.*

VI. City Square Mobility Model

In this section, we introduce a new mobility model for pedestrian movement in city parks or squares. The floating content concept can be utilized in such places, as there are typically a constant flow of people, monuments and other tourist attractions, and interesting shops can be found nearby.

In the square mobility model, nodes move in a rectangular area depicted in Fig. 5, bounded by roads and buildings so that the only entry (and exit) points are located at the four corners of the zone. For intra zone movement, we adapt the well known random waypoint (RWP) model [8]–[12]. Formally, the *square mobility model* is defined follows.

City Square Mobility Model:

- *Nodes arrive according to a Poisson process to four corners of the square, from where they travel immediately to a random point in the anchor zone.*
- *Upon reaching a waypoint, with probability of $P^{(exit)}$ the node chooses (uniformly in random) one of the corners as the exit waypoint. Otherwise, the node chooses uniformly in random a next waypoint within the square.*
- *Node moves at a constant velocity of v.*

Consequently, each visit in the anchor zone comprises two legs from/to a corner to/from a random waypoint, and X RWP style legs within the square, where X is a geometrically distributed random variable with parameter $P^{(\text{exit})}$, i.e.,

$$\mathrm{E}[X]=\frac{1-P^{(\text{exit})}}{P^{(\text{exit})}}.$$

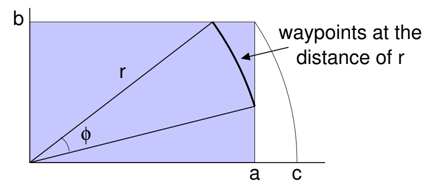

Fig. 6. Rectangular anchor zone with side lengths a and b.

We note that squares with tourist attractions may have a lower exit probability $P^{(\text{exit})}$ than squares where pedestrian mainly pass through the region.

A. Mean path length and sojourn time

Consider the $a \times b$ rectangular area depicted in Fig. 6. Without loss of generality, we can assume that $a \geq b$. Let c denote the diagonal of the rectangle, $c=\sqrt{a^2+b^2}$, and r the distance a node travels from the entry corner to the first waypoint. The equi-distant points from the corner of the rectangle lie on a arc bounded by the sides of the rectangle. We let $\phi=\phi(r)$ denote the corresponding angle. The mean distance from the corner to a random waypoint reads then

$$\mathrm{E}[L_e]=\frac{1}{ab}\int_0^c r^2\,\phi(r)\,dr. \tag{10}$$

where angle $\phi(r)$ is (see Fig. 6),

$$\phi(r)=\begin{cases}\frac{\pi}{2}, & 0\leq r\leq b,\\ \frac{\pi}{2}-\arccos(\frac{b}{r}), & b\leq r\leq a,\\ \frac{\pi}{2}-\arccos(\frac{b}{r})-\arccos(\frac{a}{r}), & a\leq r\leq c.\end{cases} \tag{11}$$

Substituting (11) into (10) then gives

$$\mathrm{E}[L_e]=\frac{\sqrt{a^2+b^2}}{3}+\frac{a^2}{6b}\ln(\frac{b+\sqrt{a^2+b^2}}{a})+\frac{b^2}{6a}\ln(\frac{a+\sqrt{a^2+b^2}}{b}),$$

and as $c=\sqrt{a^2+b^2}$, we finally obtain

$$\mathrm{E}[L_e]=\frac{c}{3}+\frac{a^2}{6b}\ln(\frac{b+c}{a})+\frac{b^2}{6a}\ln(\frac{a+c}{b}). \tag{12}$$

The mean length for the RWP style leg within the rectangular area is available from [8]:

$$\mathrm{E}[L_{\text{rwp}}]=\frac{1}{15}\left(\frac{a^3}{b^2}+\frac{b^3}{a^2}+c(3-\frac{a^2}{b^2}-\frac{b^2}{a^2})\right)+\frac{1}{6}\left(\frac{b^2}{a}\operatorname{arcosh}\frac{c}{b}+\frac{a^2}{b}\operatorname{arcosh}\frac{c}{a}\right). \tag{13}$$

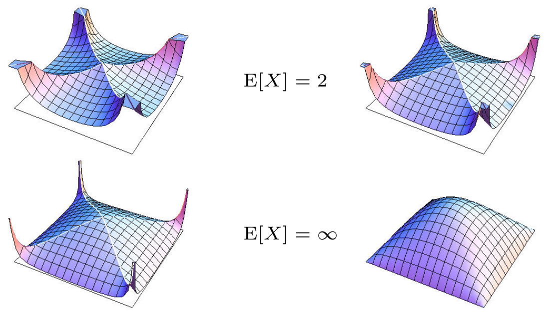

Fig. 7. Stationary spatial node distribution in unit square with the mean number of RWP legs $\mathrm{E}[X] = 0, 2, 10$ and ∞.

Finally, the *average distance a node travels* within the anchor zone reads,

$$\mathrm{E}[L] = 2\,\mathrm{E}[L_e] + \frac{1 - P^{(\mathrm{exit})}}{P^{(\mathrm{exit})}} \mathrm{E}[L_{\mathrm{rwp}}], \qquad (14)$$

where $\mathrm{E}[L_e]$ and $\mathrm{E}[L_{\mathrm{rwp}}]$ are given by (12) and (13), respectively. This is an important quantity as it gives the *mean sojourn time* of a node,

$$\mathrm{E}[T] = \mathrm{E}[L]/v. \qquad (15)$$

and the *mean number of nodes* in the area (cf. Little's result),

$$N = \lambda\,\mathrm{E}[L]/v,$$

where velocity v was some constant and λ the arrival rate. For a random leg or visit specific velocity we have $\mathrm{E}[T] = \mathrm{E}[L]\,\mathrm{E}[1/v]$ and $N = \lambda\,\mathrm{E}[L]\,\mathrm{E}[1/v]$.

B. Contact Rate

Two nodes are said to encounter each other when the distance between them becomes less than the transmission range d (i.e., Gilbert's model [13]). The total contact rate, denoted by R, is the total rate at which nodes meet in the anchor zone. It is one important quantity when assessing the feasibility of opportunistic schemes such as the floating content. The total contact rate can be computed exactly for the square mobility model using the formulæ given in [2]. For simplicity, here we simply estimate it by assuming that the movement pattern in the anchor zone were isotropic (all directions equally likely). Additionally, we assume point contacts, i.e., typical transitions are much longer than the transmission range d.

In this case, similarly as in Section V-B, the contact rate per unit distance for a single node is [2] $\tilde{\lambda} = 8nd/\pi$, and hence the contact rate per unit time is

$$\lambda_c = \frac{8}{\pi} n d v.$$

Node density is $n = N/A$, where A is the area of the anchor zone, $A = ab$, and the total contact rate is

$$R = \frac{N \cdot \lambda_c}{2} = \frac{4}{\pi}\,\frac{dvN^2}{A}. \qquad (16)$$

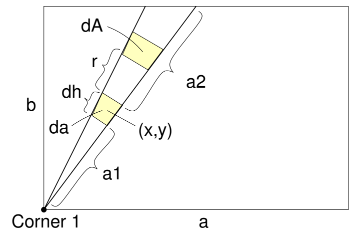

Fig. 8. Node density of enter/exit legs for square mobility model.

C. Starionary spatial node density

One important quantity is the spatial node density distribution. The corresponding results for RWP are given in [11],

$$f_{\mathrm{rwp}}(\mathbf{r}) = \frac{1}{A^2\,\mathrm{E}[L_{\mathrm{rwp}}]} \int_0^{\pi} a_1 a_2 (a_1 + a_2)\, d\theta,$$

where a_1 and a_2 are the distances to the boundary from $\mathbf{r}$ in directions θ and $\theta + \pi$, respectively. In our model, the mean number of nodes moving on RWP legs is $N_{\mathrm{rwp}} = \lambda\,\mathrm{E}[X]\,\mathrm{E}[L_{\mathrm{rwp}}]/v$, and the density of such nodes is

$$n_{\mathrm{rwp}}(\mathbf{r}) = N_{\mathrm{rwp}}\, f_{\mathrm{rwp}}(\mathbf{r}) = \frac{\lambda\,\mathrm{E}[X]}{A^2 v} \int_0^{\pi} a_1 a_2 (a_1 + a_2)\, d\theta.$$

Additionally, we have entry/exit legs occuring at rate 2λ. Consider first a node arriving/exiting via Corner 1 as illustrated in Fig. 8. The sample node moves through a differential area da at (x, y) if its other waypoint is in dA, where

$$da = (a_1 + r) \cdot d\theta \cdot dr, \qquad dA = a_1 \cdot d\theta \cdot dh.$$

The rate of such transitions is $(2\lambda/4) \cdot dA/(ab)$. These nodes spent time dh/v inside da, and according to Little's result, the mean number of them in da is $(\lambda/2) \cdot dA/(ab) \cdot dh/v$. Therefore, their contribution to the node density at da is

$$dn_1 = \frac{\lambda \cdot dA \cdot dh}{2abv \cdot da} = \frac{\lambda(a_1 + r)}{2abva_1}\, dr,$$

and integrating over r gives

$$n_1(x, y) = \frac{\lambda(2 + a_2/a_1)a_2}{4abv} \qquad (17)$$

where $a_1 = \sqrt{x^2 + y^2}$ and $a_2 = \sqrt{x^2 + y^2}(\min\{a/x, b/y\} - 1)$. Due to the symmetry in rectangular topology, the density of nodes entering or exiting the area is

$$n_e = n_1(x, y) + n_1(a - x, y) + n_1(x, b - y) + n_1(a - x, b - y),$$

and the total node density is the sum $n_e(\mathbf{r}) + n_{\mathrm{rwp}}(\mathbf{r})$. Fig. 7 illustrates how the shape of the node density distribution changes as a function of $\mathrm{E}[X]$.

VII. Floating Content - Case Study

In this section, we will present a case study of floating content using real world mobility data collected from three different cities, namely *Abbey Road Crossing*,[1] London, *Wrigley Field Crossing*,[2] Chicago, Illinois, and *Place Centrale*

[1]http://www.abbeyroad.com/crossing/
[2]http://www.earthcam.com/usa/illinois/chicago/wrigleyfield/

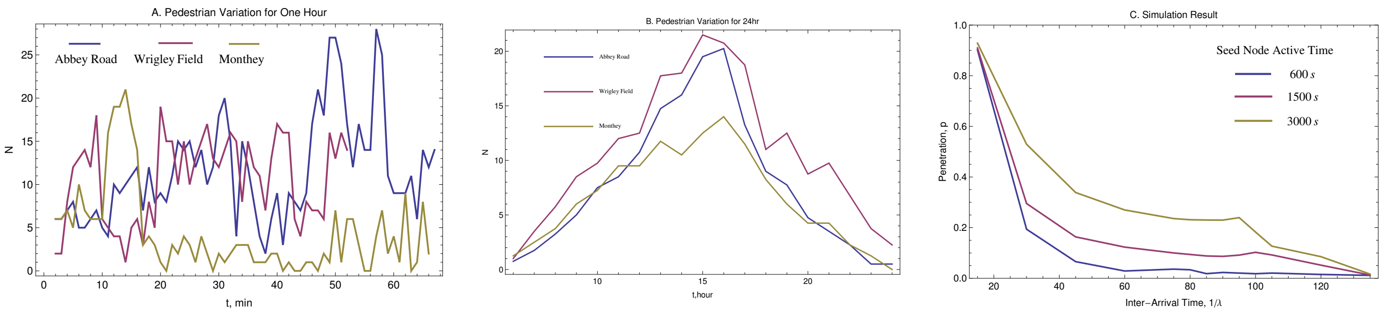

Fig. 9. A) Number of Pedestrian, N for One Hour Interval B) Number of Pedestrian, N for 24 Hour Interval C) Simulation Result: Content penetration for the example in VII-A

de Monthey nord en directe,[3] Switzerland. By using freely available real time webcams from the internet, we were able to gather a small scale pedestrian data by observing the number of pedestrians at regular time intervals.[4] The data collected is illustrated in Fig. 9a. In order to capture the short term variations in the number of nodes, we took sample frames in one minute interval from the video streams starting midday for approximately one hour, and counted the number of pedestrians in each frame. The mean number of nodes was 11.6, 10.6 and 4.4 for *Abbey Road Crossing*, *Wrigley Field Crossing* and *Place Centrale de Monthey nord en directe*, respectively. Fig. 9b gives the average number of pedestrian variation observed in a 24-hour cycle.

Next we limit ourselves to consider the *Abbey Road Crossing*. The data from it corresponds to an urban region of approximately $A{=}1218\,\mathrm{m}^2$ area with side lengths of $29\,\mathrm{m} \times 42\,\mathrm{m}$. Correspondingly, we choose $a{=}29\,\mathrm{m}$ and $b{=}42\,\mathrm{m}$ in the square mobility model. With these values,

$$\mathrm{E}[L_e] \approx 27.44\,\mathrm{m}, \quad \text{and} \quad \mathrm{E}[L_{\mathrm{rwp}}] \approx 18.67\,\mathrm{m}.$$

Assuming further that all nodes have a transmission range of $d = 10\,\mathrm{m}$ and travel at a constant speed of $v = 1\,\mathrm{m/s}$, then the average sojourn time is

$$T_{sj} = \frac{2\,\mathrm{E}[L_e] + \frac{1-P^{(\mathrm{exit})}}{P^{(\mathrm{exit})}}\,\mathrm{E}[L_{\mathrm{rwp}}]}{v} \approx 36.21\,\mathrm{s} + \frac{18.67\,\mathrm{s}}{P^{(\mathrm{exit})}}, \quad (18)$$

while (16) gives,

$$R = \frac{\nu N^2}{2} \;\Leftrightarrow\; \nu = \frac{2R}{N^2} = \frac{8dv}{\pi A} = \frac{40}{609\pi} \approx 0.0209. \quad (19)$$

Finally, applying the necessary condition in (3) yields,

$$N > \frac{47.83\,P^{(\mathrm{exit})}}{(1-p^*)(18.67 + 36.21\,P^{(\mathrm{exit})})}. \quad (20)$$

Fig. 10 depicts the necessary condition (20) for the system to stabilize at p for $P^{(\mathrm{exit})} = 1.0, 0.75, 0.5, 0.25$ (from top to bottom). That is, the region above the curves is where floating content is feasible. Comparing to our real-world data, we see that the mean number of pedestrians for the three squares is well above the curves in Fig. 10, suggesting that content sharing in urban locations without an infrastructure is feasible.

[3]http://www.idelec.ch/webcam/webcam_monthey.htm

[4]How well the actual movement follows the Square Mobility Model, e.g., in terms of the number of turns, needs further analysis.

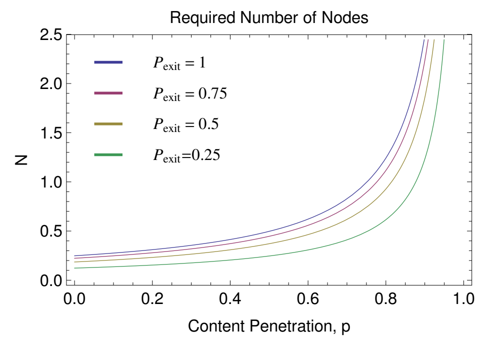

Fig. 10. Required number of nodes to float content for some values of $P^{(\mathrm{exit})}$.

A. Example

Consider a square anchor zone with sides $a{=}b{=}74\,\mathrm{m}$ and exit probability $P^{(\mathrm{exit})}{=}0.4355$. Then, (14) gives $\mathrm{E}[L]{=}163.262\,\mathrm{m}$ and (16) gives $R{=}0.00258N^2$. Consequently, the criticality condition in (3) will be $1/\lambda{<}75.4\,\mathrm{s}$.

Fig. 9.c depicts the corresponding simulation results obtained with the ONE simulator [14]. The system is initialized with the seeder node staying in the area (active time) for $600\,\mathrm{s}$, $1500\,\mathrm{s}$ and $3000\,\mathrm{s}$ after which the content is left to float in the area on its own. We observe that a higher penetration of content is achieved when the seeder stays longer (active time), in effect, passing the content to more incoming nodes. This prevents the situations where a single carrier walks out of the anchor zone and the content disappears.

Another point to note from the simulation result is that operating the system above the criticality threshold ($1/\lambda{=}75.4$) increases the proportion of nodes having the content, p, thereby guaranteeing content *availability*.

VIII. Discussion

In our previous work [15], we showed that acquisition times for content upon a node entering an anchor zone are typically a few minutes. In [15] we also presented example applications and we now contrast them against the results in this paper, in particular the square mobility case.

One of the main findings in this paper is that content availability is substantially improved by having the content publisher remain in the anchor zone for a possibly long duration after publishing. Note that this does not mean the publisher would have to remain immobile.

For applications where users are on the move (e.g., digital graffiti or regional chat from [15]), the requirement to stay in the anchor zone would be too onerous. However, given that these applications attempt to establish communications between two or more users, you would likely use a larger anchor zone, i.e., effectively meeting the sojourn constraints.

A small-scale example is the auction [15] in a flea market. The area is small, but as people are likely to remain browsing the market for long durations, it is probable that the system meets the more stringent constraints defined in this paper.

The collaborative sensing application from [15] is somewhat different. It has highly mobile users and information with a very limited range of interest, leading to small anchor zones. However, since all users are effectively publishing the same item (e.g., signal strengths of WLAN networks to provide information about connectivity), the overall system is likely to meet the constraints for successful floating.

Summarizing our examples and results, we conclude:

- During the initial phase, it is important to create several copies of the content in order to avoid an early extinction. This applies to both small and large systems.
- In a small system nodes have to meet several new nodes before departing. Large-scale floating systems are not prone to stochastic fluctuations and it is sufficient that each node meets on average "one" new node [2].
- The square mobility model extends the analytical evaluation of floating content into city squares and our real-world data from three squares confirms that typically they meet the constraints for floating.
- Although small squares put more stringent requirements on how content publishers should behave, we believe that in real-world scenarios these are not limiting factors but that they would mostly conform to normal user behavior.

IX. Conclusion

Past work on floating content has focused on studying large floating content systems' capability to store content "permanently" in the anchor zone. In particular, the criticality condition tells us under what circumstances the content avoids extinction, given it has been first distributed sufficiently well in the anchor zone. In contrast, in this paper, we have studied (i) the information availability (will a random node acquire it) and (ii) the initial *bootstrapping* phase when a new content is created (does the content disappear soon after the creator departs or not). By analyzing different (stochastic) models capturing the essential characteristics of the system, we found that unlike very large systems, smaller systems can be very prone to stochastic fluctuations and in order to avoid the aforementioned pitfalls, such systems should be operating clearly above the criticality threshold.

Additionally, we have considered human mobility in an open city square, where pedestrians arrive, move around, and eventually leave. Such an environment is seen as one prospective scenario for the floating content scheme. To better understand such scenario, we first defined a sound square mobility model with one free parameter, $P^{(\text{exit})}$, that controls the shape of the path nodes take in the square before departing. Asymptotically, when $P^{(\text{exit})} \to 0$ one obtains the RWP model, and when $P^{(\text{exit})} = 1$ the nodes make one stop and then leave the area. We have given several analytical results for the square mobility model that facilitate, e.g., adjusting the model parameters for simulation studies. Further, we have carried out a mobility measurement by using freely available webcams in the internet offering video streams from squares around the world. Based on the analytical and numerical results, it appears that people visiting such squares could support floating content type of opportunistic networking scheme.

Our future work includes more detailed analysis of both the initial content distribution phase and the square mobility model. It would be also interesting to see if mobility parameters could be estimated automatically from a suitable video stream, thus enabling analysis of very long time periods.

References

[1] J. Kangasharju, J. Ott, and O. Karkulahti, "Floating content: Information availability in urban environments," in *Proceedings of IEEE PerCom Workshops*, 2010, pp. 804–808.

[2] 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 *IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 3123–3131.

[3] J. Ott, E. Hyytiä, P. Lassila, T. Vaegs, and J. Kangasharju, "Floating Content: Information Sharing in Urban Areas," in *Proc. of IEEE Percom 2011*, March 2011.

[4] A. Villalba Castro, G. Di Marzo Serugendo, and D. Konstantas, "Hovering information: Self-organizing information that finds its own storage," School of Computer Science and Information Systems, Birkbeck College, London, UK, Tech. Rep. BBKCS707, Nov. 2007.

[5] O. Helgason, E. Yavuz, S. Kouyoumdjieva, L. Pajevic, and G. Karlsson, "A mobile peer-to-peer system for opportunistic content-centric networking," in *ACM/SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds (MobiHeld)*, Aug. 2010.

[6] V. Lenders, M. May, G. Karlsson, and C. Wacha, "Wireless ad hoc podcasting," *ACM/SIGMOBILE Mobile Computing and Communications Review*, Jan. 2008.

[7] N. Thompson, R. Crepaldi, and R. Kravets, "Locus: A location-based data overlay for disruption-tolerant networks," in *Workshop on Challenged Networks*, Chicago, IL, Sep. 2010.

[8] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa, "Stochastic properties of the random waypoint mobility model," *ACM/Kluwer Wireless Networks*, vol. 10, no. 5, Sep. 2004.

[9] W. Navidi and T. Camp, "Stationary distributions for the random waypoint mobility model," *IEEE Transactions on Mobile Computing*, vol. 3, no. 1, pp. 99–108, January-March 2004.

[10] J. L. Boudec, "On the stationary distribution of speed and location of random waypoint," *IEEE Transactions on Mobile Computing*, Dec. 2005.

[11] E. Hyytiä, P. Lassila, and J. Virtamo, "Spatial node distribution of the random waypoint mobility model with applications," *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 680–694, Jun. 2006.

[12] E. Hyytiä and J. Virtamo, "Random waypoint mobility model in cellular networks," *Wireless Networks*, vol. 13, no. 2, Apr. 2007.

[13] E. Gilbert, *Random Plane Networks*. J.SIAM, December 1961, vol. 9, no.4.

[14] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE Simulator for DTN Protocol Evaluation," in *SIMUTools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*. New York, NY, USA: ICST, 2009.

[15] J. Ott and J. Kangasharju, "Opportunistic content sharing applications," in *Proceedings ACM MobiHoc Workshop on Emerging Name-Oriented Mobile Networking Design-Architecture, Algorithms, and Applications*, Jun. 2012, pp. 19–24.