

When Does Content Float? Characterizing Availability of Anchored Information in Opportunistic Content Sharing

Esa Hyytiä[†], Jorma Virtamo[†], Pasi Lassila[†], Jussi Kangasharju[‡], Jörg Ott[†]

Aalto University School of Science[†]
and Technology, Finland

Helsinki Institute for Information Technology[‡]
Helsinki University, Finland

Abstract—We consider an opportunistic content sharing system designed to store and distribute local spatio-temporal “floating” information in uncoordinated P2P fashion relying solely on the mobile nodes passing through the area of interest, referred to as the *anchor zone*. Nodes within the anchor zone exchange the information in opportunistic manner, i.e., whenever two nodes come within each others’ transmission range. Outside the anchor zone, the nodes are free to delete the information, since it is deemed relevant only for the nodes residing inside the anchor zone. Due to the random nature of the operation, there are no guarantees, e.g., for the information availability. By means of analytical models, we show that such a system, without any supporting infrastructure, can be a viable and surprisingly reliable option for content sharing as long as a certain criterion, referred to as the *criticality condition*, is met. The important quantity is the average number of encounters a randomly chosen node experiences during its sojourn time in the anchor zone, which again depends on the communication range and the mobility pattern. The theoretical studies are complemented with simulation experiments with various mobility models showing good agreement with the analytical results.

I. INTRODUCTION

Content sharing via the Internet is a widespread means for people to foster their relationships irrespective of physical distance, cf. the phenomenal popularity of Facebook and Twitter. While network-based social applications are essential to overcome distances and connect people around the world, relying on infrastructure for location-aware services may often be undesirable [1], [2], [3]. We propose and analyze a fully distributed variant of an ephemeral content sharing service, solely dependent on the mobile devices in the vicinity using principles of opportunistic networking. The net result is a best effort service for *floating content* in which: 1) information dissemination is geographically limited; 2) the lifetime and spreading of information depend on interested nodes being available; 3) traffic can only be created and caused locally; and 4) content can only be added, but not deleted.

We analyze fundamental properties of the floating content concept [1]. The system is assumed to consist of independent users, each moving according to some mobility pattern. Moreover, there is a specific region, the so-called *anchor zone*,

where users exchange the information item on condition that they are within each others’ transmission ranges. Once a user exits the anchor zone, she can forget the information that is considered obsolete outside the anchor zone. The transmission range is assumed to be fixed, same for all users, and small relative to the dimensions of the anchor zone. No assumption is made regarding whether the number of nodes in the region is large enough to guarantee instantaneous connectivity, i.e., the network does not need to percolate. Thus, user mobility is the key factor through which information is propagated.

The floating content concept is a specific example of networks with intermittent connectivity or so-called delay tolerant networks (DTN). While the performance limits of wireless multihop networks assuming that the users form a connected network (i.e., the network percolates) have been analyzed extensively in the past years, see, e.g., [4], [5], corresponding analytical results for DTNs are much more scarce. Much of the research on DTNs has focused on protocol and algorithm development based on simulations. Recently, analytical results have also started to emerge. The delay of the information dissemination for epidemic-style routing algorithms assuming exponential inter-meeting times has been considered in [6], [7]. The validity of the exponential inter-meeting time assumption for particular mobility models (synthetic models and real traces) has been studied in [6], [8]. The scaling properties of the delay for any opportunistic routing scheme below the percolation threshold have been analyzed in [9] and upper bounds have been derived in [10]. Optimal policies for routing in DTNs with different objective functions and constraints have been recently considered, e.g., in [11], [12]. However, the available results do not capture the essential aspect of the floating content application, where we need to characterize the notion of information availability in the system.

We develop models for analyzing the feasibility of the floating content concept by using spatial modeling techniques adopted from physics [13], similarly as have been used for modeling the capacity of so-called dense wireless multihop networks, see, e.g., [14], [15]. The fundamental objective is to establish the criticality condition for the system, i.e., an explicit relation which determines when the information remains available within the anchor zone infinitely long at the fluid limit (large population). This obviously depends on

[†]This work was supported by TEKES as part of the Future Internet program of TIVIT (Finnish Strategic Centre for Science, Technology and Innovation in the field of ICT), and by Teknologiateollisuus ry within REDI project.

various mobility specific parameters, and we show that, in particular, the node encounter rate in the anchor zone and the node arrival rate are the key factors to this end.

More specifically, in Section II, we introduce the floating content concept. Section III gives generic results characterizing the node encounter rates, i.e., rate at which nodes come into contact with each other, and the mean contact times via the notion of the angular node flux.

Then we consider three mobility scenarios. First, in Section IV, in a non-spatial approach, users are assumed to spend a considerable time inside the anchor zone so that the information can be assumed to be homogeneously distributed in the zone, and we derive an ordinary differential equation for the rate of change in the information availability. In this case, the criticality condition of this fluid model turns out to have a particularly intuitive form, $2\mathcal{R} > \Lambda$, where \mathcal{R} denotes the total rate of node encounters in the anchor zone and Λ is the node arrival rate to the anchor zone. Note that $2\mathcal{R}/\Lambda$ is the mean number of encounters a randomly chosen node experiences during its visit. This approach also allows us to estimate the lifetime of the corresponding stochastic system.

In Section V, the criticality condition is further examined with elementary mobility models, for which we first derive exact expressions for encounter rates. Then, the non-spatial criticality condition (see above) is applied to estimate the actual criticality condition for each mobility pattern. The approximation in this hybrid approach is that the spatial aspect of the information exchange is not modelled exactly.

Finally, in Section VI, we assume that users pass through the area along a straight line. In this case, the spatial aspects are taken exactly into account, and an integral equation is derived characterizing the probability of acquiring the information at a given distance from the center. Unfortunately, the criticality condition cannot be solved analytically. However, numerical analysis yields a simple explicit criticality condition. Also auxiliary parameters can be derived, such as the mean length a user moves before acquiring the information (if at all).

The analytical results are supported by extensive numerical results obtained with a special purpose simulator written in C. Notably, the results show a good match with the behavior predicted by the more abstract models, and indeed, the floating content scheme is applicable only when the criticality condition is met.

II. FLOATING CONTENT CONCEPT

Next we briefly outline the main concepts and operation of a floating content system. More details can be found in [1].

A. System Basics

We assume that all users are mobile nodes and that there is no supporting infrastructure for the system.¹ The users are interested in information items “posted” by other users. They use mobile phones or similar devices to communicate, so that ample storage capacity is available of which, say, 1 GB is

¹Content dissemination from fixed access points would lead us towards the PodNet model [16] paired with geographically limited distribution.

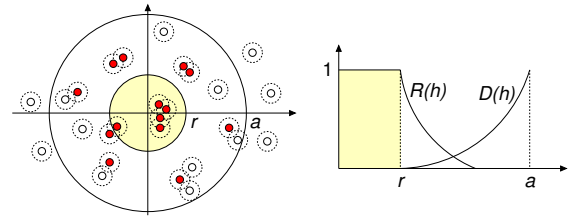


Fig. 1. An anchor zone of an item, mobile nodes and their communication ranges: the content item gets replicated across and deleted from nodes as a function of the distance from the anchor point. The fraction of nodes carrying an item (solid nodes) tends to 1 inside the anchor zone r and decreases until, after an availability threshold a , no more copies are found. The replication and deletion probabilities, $R(h)$ and $D(h)$, are illustrated on right.

used for floating content. The devices have wireless interfaces (e.g., Bluetooth or WLAN) for ad-hoc communication within a certain range. The nodes cooperate by replicating content among interested parties as we describe below.

Each item has an anchor zone, which is a real world area in which the items should be made available. We assume circular anchor zones defined by a center point and a radius.² Fig. 1 shows an example of an anchor zone and nodes.

As noted above, interested nodes keep copies of information items floating around in the anchor zone by probabilistically replicating the items when they meet. We explicitly allow information items to disappear from the system and provide no guarantees about their availability. If no (or too few) nodes are around to replicate an information item, the corresponding information items will disappear (over time). Content items may be tagged with a lifetime and are discarded thereafter.

Anchor zones require nodes to be able to determine their position, e.g., by using GPS receivers or triangulation-based methods using WLAN access points, cellular base stations, etc. Since the system is probabilistic, there are no strict requirements on the accuracy of positioning techniques; nodes are only required to agree on measurement parameters and the overall operation to determine the extent of anchor zones.

B. System Operation

A node generates an information item I and assigns an anchor zone (defined by its center and its radius), as well as, a (tentative) time to live for this item. We require that the generating node is within the zone when an item is created.

If 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 item I to B . Since replication is based purely on the location of nodes, in a simple case, every node in the anchor zone should have a copy of the item. Nodes leaving the anchor zone are free to delete their copy of the item.

In practice, the replication and deletion works as follows. Consider a node A having an item I , with an anchor zone defined by center point c and radius r . Let h denote the distance of node A from point c . When node A meets another node B , A will replicate item I to B with probability of $R(h)$,

²Any other shape is also possible, provided that it can be expressed in a relatively succinct manner since it is communicated along with the item I .

where the replication probability $R(h) \in [0, 1]$ is 1 inside the anchor zone, and decreases to 0 outside the zone (see Fig. 1).

For deletion of items, we can define a similar function $D(h)$ which is 0 if the node A is in the anchor zone and some (increasing) function when A leaves the anchor zone. Allowing the item to survive outside the anchor zone provides additional protection against items disappearing when nodes move outside the anchor zone for a brief moment and then return. We define an availability distance a (see Fig. 1) beyond which copies are deleted. The deletion function essentially serves the purpose of prioritization when the buffer gets full. It is evaluated upon each encounter with another node or whenever there is a need to discard content items to free buffer space. Further storage management issues, e.g., limiting effects of spam, are covered in [1].

In our evaluation presented below, we use a step function: $D(h) = 1$ for $h > r$ and, $R(h) = 0$ for $h > r$, i.e., the content is unconditionally deleted outside the anchor zone, and replication occurs only inside the anchor zone.

III. PRELIMINARIES

A. Fluxes and stationary distribution

Consider an arbitrary mobility model where n mobile users, referred to as nodes, move according to some pattern in a subset \mathcal{A} of the plane. In order to analyze the floating content concept, it is important to characterize the movement, e.g., how often a node crosses a given line segment. To this end, let $\varphi(\mathbf{r}, \phi)$ denote the so-called angular node flux at \mathbf{r} in direction ϕ , which is defined as the rate of nodes moving in direction $(\phi, \phi + d\phi)$ across a small perpendicular line segment of ds divided by $ds \cdot d\phi$ at the limit $ds, d\phi \rightarrow 0$ [13], [15], as illustrated in Fig. 2. Similarly, $\Phi(\mathbf{r})$ denotes the so-called scalar node flux at \mathbf{r} , which is the sum of angular fluxes:

$$\Phi(\mathbf{r}) := \int_0^{2\pi} \varphi(\mathbf{r}, \phi) d\phi.$$

Let $\mathbf{r}(s)$ denote a parametric representation of an arbitrary curve \mathcal{S} with s denoting the distance along the curve from one end towards the other. The total flux of nodes from left to right across \mathcal{S} is [17]

$$\lambda(\mathcal{S}) = \int_{\mathcal{S}} \int_0^{\pi} \sin \phi \cdot \varphi(\mathbf{r}(s), \theta(s) + \phi) d\phi ds, \quad (1)$$

where $\theta(s)$ denotes the direction of the tangent at $\mathbf{r}(s)$. From (1) it follows that the arrival rate $\lambda = \lambda(\mathbf{r}, d)$ to a small disk with radius d at \mathbf{r} is

$$\lambda(\mathbf{r}, d) = 2d \cdot \Phi(\mathbf{r}).$$

Let $n(\mathbf{r})$ denote the node density at \mathbf{r} so that the stationary node distribution $f(\mathbf{r}) = n(\mathbf{r})/n$. Then,

$$n(\mathbf{r}) \cdot v(\mathbf{r}) = \Phi(\mathbf{r}) \Leftrightarrow f(\mathbf{r}) = \frac{\Phi(\mathbf{r})}{n \cdot v(\mathbf{r})}, \quad (2)$$

where $v(\mathbf{r})$ is the mean velocity of a node at \mathbf{r} . These hold for an arbitrary movement pattern.

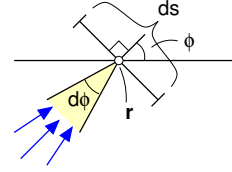


Fig. 2. Angular node flux $\varphi(\mathbf{r}, \phi)$ is a rate of nodes crossing a perpendicular line segment ds moving in direction $(\phi, \phi + d\phi)$ at the limit $ds, d\phi \rightarrow 0$.

B. Encounter rates and contact times

1) *Encounter rate \mathcal{R}* : Consider next two nodes A and B moving independently according to some mobility model(s). For simplicity, we further assume a constant velocity v . Let $A(t)$ and $B(t)$ denote their positions at time t , respectively. Define $C(t) = A(t) - B(t)$, referred to as the *difference walker* in [18], i.e., $C(t)$ defines a certain mobility model. In particular, events when $C(t)$ visits near the origin correspond to encounters between nodes A and B , i.e., in some sense “collide”. We assume a constant transmission range d and an encounter or a “collision” occurs when the distance between two nodes decreases below d . Furthermore, let $f_c(\mathbf{r})$ denote the stationary distribution of node C , and v_c the mean velocity of node C at the origin. The quantity v_c is the *mean relative encounter velocity* of nodes A and B , and thus depends on the mobility model and the area. The encounter rate of the two nodes, in the limit $d \rightarrow 0$, is

$$v_{ab}(d) = 2d \cdot \Phi_c(0) = 2d \cdot v_c \cdot f_c(0) = 2d \cdot v_c \int_{\mathcal{A}} f(\mathbf{r})^2 d^2\mathbf{r}.$$

For a large number of nodes N , there are approximately $N^2/2$ pairs, and the total rate of encounters is

$$\mathcal{R}(d) = d \cdot v_c \int_{\mathcal{A}} n(\mathbf{r})^2 d^2\mathbf{r}. \quad (3)$$

It follows that the mean encounter velocity about \mathbf{r} is³

$$v_c(\mathbf{r}) = \frac{2v}{\Phi(\mathbf{r})^2} \int_0^{2\pi} \int_0^{2\pi} \varphi(\mathbf{r}, \phi) \cdot \varphi(\mathbf{r}, \phi + \theta) \cdot \sin(\theta/2) d\theta d\phi,$$

which then gives

$$v_c = \frac{\int n(\mathbf{r})^2 \cdot v_c(\mathbf{r}) d^2\mathbf{r}}{\int n(\mathbf{r})^2 d^2\mathbf{r}}.$$

Substituting the above into (3), and using (2), one obtains

$$\begin{aligned} \mathcal{R}(d) &= d \int_{\mathcal{A}} n(\mathbf{r})^2 \cdot v_c(\mathbf{r}) d^2\mathbf{r} \\ &= \frac{2d}{v} \int_{\mathcal{A}} \int_0^{2\pi} \int_0^{2\pi} \varphi(\mathbf{r}, \phi) \varphi(\mathbf{r}, \phi + \theta) \sin \frac{\theta}{2} d\theta d\phi d^2\mathbf{r}. \end{aligned} \quad (4)$$

Eq. (4) provides also the density function for encounter angles θ , i.e. the rate at which nodes “collide” with angle θ . As an

³Note that $\sqrt{1 - \cos \theta} = \sqrt{2} |\sin(\theta/2)|$.

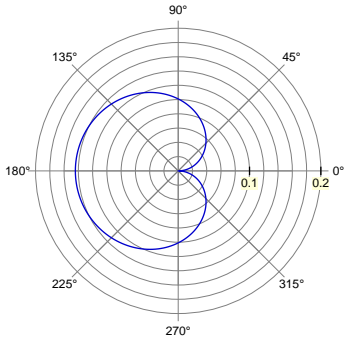


Fig. 3. Angular encounter rate, rate of encounters with angle $(\theta, \theta + d\theta)$ divided by $d\theta$ at the limit $d\theta \rightarrow 0$, is close to $\sin(\theta/2)$.

example, Fig. 3 illustrates this for standard RWP model in unit disk with a constant velocity. It turns out that the form of angular encounter rate is very close to $\sin(\theta/2)$, $\theta \in [0, 2\pi)$.

2) *Contact time duration*: Contact time duration (intermeeting time in [18]) refers to the time duration two nodes remain within the transmission range d of each other. Only when the contact time duration is sufficiently long the content can be transmitted successfully. The mean contact time is (cf. Little's result for queueing systems [19])

$$T(d) = \frac{\pi d^2 \cdot f_c(0)}{\mathcal{R}(d)} = \frac{\pi \cdot d}{2 \cdot v_c}, \quad (5)$$

which holds for an arbitrary independent mobility model when d is small. Alternatively, observing first that the nodes travel straight through a small disk, means that the mean distance is equal to the mean height, which is the area divided by the diameter, $\pi d^2 / (2d) = \pi d / 2$, yielding (5).

IV. ANALYSIS OF NON-SPATIAL MODEL

Here we regard the R -disk representing the anchor zone just as a locale (local center, building, room) where the nodes enter and then spend some time and finally exit. We assume that the nodes move randomly within the locale and the time spent there is relatively long so that the exact points of entrance and exit do not have any bearing. The whole population is assumed to be large and "well mixed" so that the number of nodes N and the proportion of the tagged nodes, i.e., users having the information, are constant. In Section IV-A, we establish the criticality condition at the fluid limit, and in Section IV-B, we describe a finite system model where, due to stochastic fluctuations, the information eventually always disappears.

A. Criticality Condition

During the sojourn time, a node encounters randomly other nodes. Consider a system consisting of solely of two nodes staying permanently in the locale and denote by ν the frequency at which they come in contact with each other, i.e., within the information exchange distance. If the total population of nodes is N , then there are $\frac{1}{2}N(N-1) \approx \frac{1}{2}N^2$ pairs and the total rate of encounters is $\mathcal{R} = \frac{1}{2}N^2\nu$. The fraction $2p(1-p)$ of these encounters are such that a non-tagged node becomes a tagged one, and the total rate of such

events is $p(1-p)N^2\nu$. This is the rate at which the size of the population of tagged nodes tends to increase. Let $1/\mu$ denote the node's mean sojourn time in the anchor zone. Then the total exit (and arrival) rate of nodes is $\Lambda = N\mu$, and the exit rate of tagged nodes is $Np\mu$, so that their net growth rate is

$$N \frac{d}{dt} p = N^2 p(1-p)\nu - Np\mu. \quad (6)$$

In equilibrium, the two terms on the right hand side are equal leading to the stationary value

$$p = 1 - \frac{\mu}{\nu N}. \quad (7)$$

The existence of a positive solution requires that the second term is less than one, i.e.,

$$N\nu/\mu > 1 \Leftrightarrow 2\mathcal{R} > \Lambda. \quad (8)$$

This is the *criticality condition* (at fluid limit). Note that the quantity on the left hand side (of the first form of the inequality) is the *average number of encounters a randomly chosen node experiences during its sojourn time*. By considering the sign of the derivative (6) one easily sees that the solution is stable; if p is greater than (7) it tends to decrease, and if it less it tends to increase. Below the criticality, the derivative is everywhere negative and the solution is driven to $p = 0$, i.e., the information cannot be sustained even in the fluid model. In addition, one has to remember, that the number of tagged nodes, $Np = N - \mu/\nu$, must be large in order to avoid accidental disappearance of the information carrying population by stochastic fluctuations.

B. Information lifetime

Let us next focus on the lifetime, i.e., how long the information will be available within the anchor zone. In contrast to the above fluid model, with a finite number of nodes, the information will be eventually lost due to the random fluctuations. Here we model the finite system by a state-dependent Markov birth-death process. Given i nodes possess the information at a given time, then the birth-death rates are

$$\lambda_i = i(N-i)\nu, \quad \text{and} \quad \mu_i = i\mu, \quad \text{for } i = 0, \dots, N. \quad (9)$$

where we assume that N is large and practically a constant. State $i = 0$ is an absorbing state in which the information is irreversibly lost. Let T_i denote the *expected lifetime* of the information on condition that initially i nodes have it. Clearly, $T_0 = 0$. Due to the memoryless property of Markov process, one can immediately write the *backward equations* for the T_i :

$$\begin{cases} T_N = q, \\ T_{N-1} = q - \frac{1}{N\mu}, \\ T_{i-1} = T_i + \frac{1}{i\mu} [i(N-i)\nu(T_i - T_{i+1}) - 1], \\ T_1 = \frac{1}{(N-1)\nu + \mu} [1 + (N-1)\nu T_2]. \end{cases} \quad (10)$$

Variable $q = T_N$ gets solved at the end. One can also write the forward equations starting from state 1, but it turns out that backward equations are numerically more stable. The same approach works for any semi-Markov birth-death process.

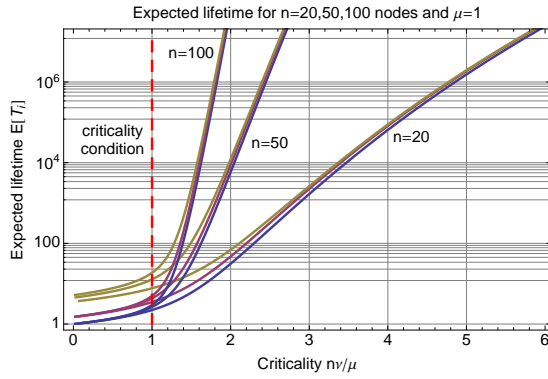


Fig. 4. Expected information lifetime with $n = 20, 50, 100$ nodes and $\mu = 1$ as a function of criticality quantity $n\nu/\mu$ in logarithmic scale. Initially $i = 1, 2, n-1$ nodes carry the information (curves from the lowest to highest).

C. Lifetime distribution

Eq. (10) gives us the mean lifetime. Additionally, one might be interested in the actual distribution. The following theorem holds for an arbitrary Markov birth-death process [20]:

Theorem 1: Let \mathbf{Q} denote the generator matrix of a continuous time birth-death process on state space $\{0, \dots, k\}$ with 0 being an absorbing state, and $\lambda_i^* > 0$, $i = 1, \dots, k-1$, and $\mu_i^* > 0$, $i = 1, \dots, k$. Suppose that the process starts at state k . Then the time to absorption is distributed as a sum of k independent exponential random variables whose rate parameters are the k non-zero eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$ of $-\mathbf{Q}$. Consequently, the time to absorption is $T = T_1 + \dots + T_k$, with $T_i \sim \text{Exp}(\lambda_i)$, and

$$E[T] = \sum_i \frac{1}{\lambda_i}, \quad \text{and} \quad \sigma^2 = \sum_i \frac{1}{\lambda_i^2}.$$

If λ_1 is considerably smaller than the other, then it *dominates* and $T \approx T_1$, which gives for the standard deviation of T $\sigma \approx E[T]$. On the other hand, standard deviation is smallest when all λ_i are equal, yielding $\sigma \approx E[T]/\sqrt{k}$.

D. Examples

Assume $N = 20, 50, 100$ nodes and $\mu = 1$. Encounter rate ν is varied so that the criticality quantity $N\nu/\mu$ obtains sensible values. Fig. 4 illustrates the resulting conditional expected lifetimes. Three curves for each N correspond to $i = 1, 2, N-1$ nodes initially having the information. We note that after the criticality threshold the expected lifetime skyrockets independently of the initial state of the system.

Consider next a birth-death process according to (9). Numerically, one notices that the ratio between the second smallest and the smallest eigenvalues λ_2 and λ_1 is high and grows extremely fast, at least for $N = 3, \dots, 7$. Thus, the information lifetime appears to be approximately exponentially distributed with parameter λ_1 . Computing eigenvalues for an arbitrary number of nodes N becomes soon numerically very challenging when N increases. However, the recursive equations (10) provide a straightforward and robust way to obtain the interesting parameter λ_1 , $\lambda_1 \approx 1/T_i$ (and $T_i \approx T_j$).

E. Remark: lifetime with a large transmission range

Here we give a brief remark about impact of a large transmission range, i.e., when the anchor zone is considerably smaller than the transmission range. In this case, the information remains (immediately) available as long as at least one user remains in the anchor zone. Assuming a Poisson arrival process with rate Λ , this system is closely related to the well-known M/G/ ∞ -queue, for which insensitivity results exist. For example, if a node arriving to an empty anchor zone creates an information, a standard busy period analysis gives the mean lifetime of the information: $E[T] = (e^{\Lambda/\mu} - 1)/\Lambda$, where $1/\mu$ denotes the node's mean sojourn time in the anchor zone.

V. ANALYSIS OF MOBILITY MODELS

In this section, we analyze the floating content concept by deriving exact expressions for the encounter rates \mathcal{R} for several elementary mobility models. Based on these, and the results in the previous section (i.e., population is still assumed to be well mixed), the criticality conditions are then established.

A. Manhattan road network

Let us first derive expressions for the total encounter rate observed in a Manhattan like network consisting of (i) road sections and (ii) perpendicular intersections.

1) *Road section:* Consider a simple case of two-directional road section with length Δ , where nodes move with a velocity of v_1 from left to right, and with a velocity of v_2 from right to left. The node arrival rates to the road section are λ_1 and λ_2 , respectively. The two directions are further assumed to be independent and sufficiently random. Then, the total rate of encounters, denoted by \mathcal{R}_s , is given by

$$\mathcal{R}_s = \lambda_1 \lambda_2 \cdot \Delta \cdot (1/v_1 + 1/v_2) = (\lambda_1 n_2 + \lambda_2 n_1) \cdot \Delta,$$

where n_1 and n_2 denote the node (line) densities on the road section. Consequently, the density of encounters is

$$\rho = \lambda_1 n_2 + \lambda_2 n_1 = n_1 n_2 (v_1 + v_2),$$

as $\lambda_i = n_i \cdot v_i$ (cf. (2)). For a uniform case with $v_1 = v_2$ and $\lambda_1 = \lambda_2$, the total rate of encounters is

$$\mathcal{R}_s = \frac{2\lambda^2 \cdot \Delta}{v}.$$

2) *Intersection:* Consider next an arbitrary perpendicular intersection where, for simplicity, we assume that all vehicles move at a constant velocity of v , have a constant transmission range of d . Let λ_1 denote the traffic rate, e.g., from left to right, and λ_2 a perpendicular traffic rate, e.g., from bottom to up. This can be seen as a special case of (4) with $\theta = \pi/2$. Consequently, the encounter rate due to λ_1 and λ_2 in an intersection is⁴

$$\mathcal{R}_x = \frac{2\sqrt{2}d}{v} \cdot \lambda_1 \cdot \lambda_2.$$

⁴Alternatively, a geometrical analysis shows that the time window during which a perpendicularly moving node must enter the intersection in order to establish a connection with an arbitrary node has length $2\sqrt{2}d/v$.

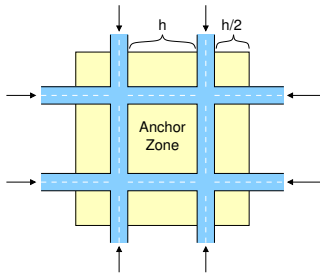


Fig. 5. In one block model, the anchor zone consists of the four road segments with length $2h$ around the block, where h is the length of side of the block. Traffic flows entering the zone are $8 \times \lambda$. Nodes move straight out of the area without making any turns.

With arbitrary velocities, $v_1, v_2 > 0$, the encounter rate due to two flows meeting in an intersection with an angle $\pi/2$ is

$$\mathcal{R}_x = 2d \cdot \sqrt{1/v_1^2 + 1/v_2^2} \cdot \lambda_1 \cdot \lambda_2, \quad (11)$$

which by conditioning then gives a general expression for a movement with an arbitrary velocity distribution. Note that there are 4 such pairs of flows in a general intersection.

3) *Example: One block model:* A single block, as illustrated in Fig. 5 represents one potential anchor zone. Let h denote the length of the block's side, so that the path around the block has length $4h$. The four road segments are further extended by $h/2$ in order to ensure arriving nodes have a fair chance to obtain the information before reaching the block. For simplicity, we further assume that nodes move at constant velocity v and do not make any turns.

The results of Section III give the total rate of encounters,

$$\mathcal{R} = 4 \cdot \frac{4\lambda^2 \cdot h}{v} + \frac{8\sqrt{2}\lambda^2 \cdot d}{v} = \frac{8\lambda^2}{v} (2h + \sqrt{2}d),$$

where $d \ll h$. Similarly, the mean number of nodes in the anchor zone is (cf., Little's result)

$$E[N] = 16 \lambda h/v.$$

Due to the simple movement model, these nodes are uniformly distributed on the 4 line segments, yielding a constant node density of $n = 2\lambda/v$. The *criticality condition*, $2\mathcal{R} > \Lambda$, gives

$$\frac{2\lambda}{v} (2h + \sqrt{2}d) = n(2h + \sqrt{2}d) > 1. \quad (12)$$

Note that the non-spatial model implicitly assumes that all nodes are well "mixed" and information can spread "freely". In Manhattan example this is not strictly the case as the information flow between the road segments is restricted. Therefore, one can assume that the criticality condition gives a somewhat optimistic estimate for the required arrival rate.

4) *Example:* Assume a block dimension of $h = 50$ m, transmission range $d = 10$ m and velocity of $v = 1$ m/s. Then the criticality condition (12) suggests that for $\lambda > 0.263$ 1/min, the mean lifetime should increase rapidly. Fig. 6 depicts the corresponding simulation results. As expected, the lifetime starts to increase fast somewhat later than (12) suggests. The discrepancy is due to the fact that the considered mobility pattern is far from the assumptions of the non-spatial model.

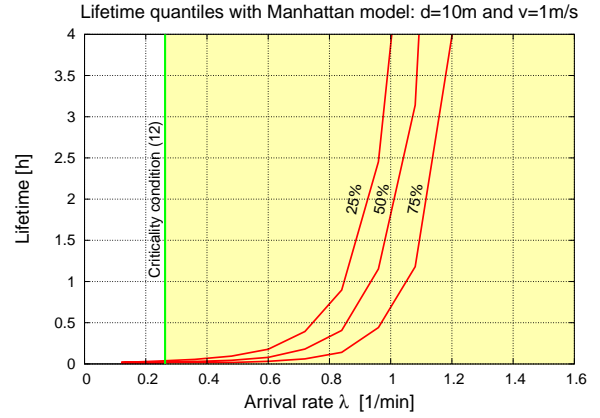


Fig. 6. Information lifetime with the Manhattan model (one block anchor zone). Three curve pairs correspond to 25%, 50% and 75% quantiles. The information is initially passed to two nodes. The vertical dashed line depicts the criticality condition (12).

B. Random Waypoint Model

RWP is a frequently used mobility model due to its simplicity facilitating the analysis, see, e.g., [21], [22], [23]. Here we simply give the necessary results and an interested reader is referred to [24]. For unit disk, the mean transition length is

$$E[L_{RWP}] = \frac{128}{45\pi}. \quad (13)$$

The angular node flux with N nodes at $\mathbf{r} = (0, r)$ is

$$\varphi(r, \phi) = \frac{45 \cdot N v}{128\pi} \cdot (1 - r^2) \sqrt{1 - r^2 \cos^2 \phi}. \quad (14)$$

Arrival rate to a concentric h -disk with radius $h < 1$ is

$$\Lambda(h) = \frac{45h(1 - h^2) \cdot N v}{64} \int_0^\pi \sin \phi \cdot \sqrt{1 - h^2 \cos^2 \phi} d\phi.$$

We note that unless $h \approx 1$, most of the nodes simply pass through the h -disk. The encounter rate within the h -disk is

$$\mathcal{R}(h, d) = \frac{45^2}{64^2 \pi} d N^2 v \int_0^h r(1 - r^2)^2 \int_0^{2\pi} \sqrt{1 - r^2 \cos^2 \phi} \int_0^{2\pi} \sqrt{1 - r^2 \cos^2(\phi + \theta)} \sin(\theta/2) d\theta d\phi dr, \quad (15)$$

i.e., for the whole unit disk one has $\mathcal{R}(d) \approx 0.573678 \cdot d N^2 v$.

Consequently, consider next the floating content concept with N nodes moving in unit disk with a constant velocity of v and having no pause times. Let the h -disk correspond to the anchor zone located at the center of the movement area. The criticality condition, $2\mathcal{R}(h, d) > \Lambda(h)$, gives

$$2d \cdot N > Q(h),$$

where

$$Q(h) = \frac{64 \pi h(1 - h^2) \int_0^\pi \sin \phi \cdot \sqrt{1 - h^2 \cos^2 \phi} d\phi}{45 \int_0^h r(1 - r^2)^2 \int_0^{2\pi} \sqrt{1 - r^2 \cos^2 \phi} \int_0^{2\pi} \sqrt{1 - r^2 \cos^2(\phi + \theta)} \sin(\theta/2) d\theta d\phi dr}$$

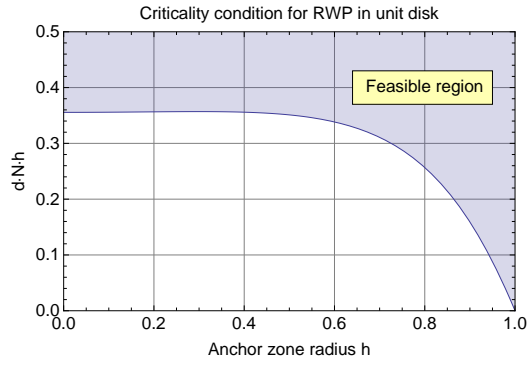


Fig. 7. Criticality condition illustrated for N RWP nodes with (small) transmission range d . Information zone radius h is on the x -axis, and y -axis corresponds to the product $dN h$. Above the curve the criticality condition holds, $2\mathcal{R}(h, d) > \Lambda(h)$. For $h = 0 \dots 0.4$, the curve is close to $16/45 \approx 0.356$, obtained at the limit $h \rightarrow 0$.

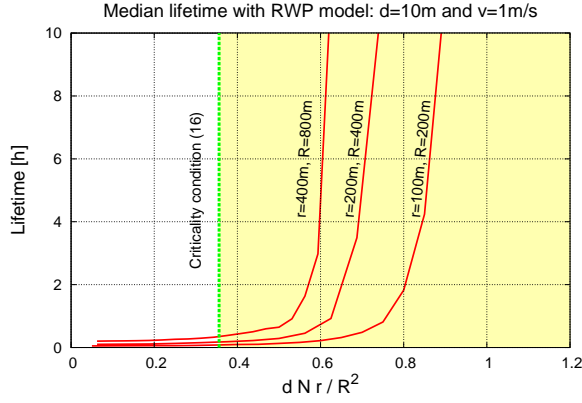


Fig. 8. Median information lifetime with the RWP model as a function of the criticality quantity dNr/R^2 , with $d = 10$ m and $R = 2r$, for three cases: (a) $R = 100$ m, (b) $R = 200$ m and (c) $R = 400$ m. The vertical dashed line corresponds to the criticality condition (16).

It turns out that $Q(h)$ behaves approximately as $1/h$ for $h < 0.4$. Criticality condition can be written as $d \cdot N \cdot h > h \cdot Q(h)/2$, and thus Fig. 7 depicts $h \cdot Q(h)/2$ as a function of zone radius h . The curve corresponds to the value of the product $dN h$ at the criticality threshold. In particular, in the limit $h \rightarrow 0$, the *criticality condition*, independently of v , becomes

$$dN h > \frac{16}{45} \approx 0.356, \quad (16)$$

which holds for a surprisingly wide range of zone radii h . A similar result has also been obtained experimentally in [1].

1) *Example:* Assume a transmission range of $d = 10$ m, a constant velocity of $v = 1$ m/s, and a movement area with radius $r = 2R$ (i.e., $h = 1/2$). The anchor zone radius is $R = 100$ m, 200 m, 400 m. In Fig. 8, on x -axis is the *criticality quantity*, dNr/R^2 , obtained by varying the number of nodes N . The y -axis depicts the median of the information lifetime. Initially, the information is passed to two nodes (solid curve) at the center of the area. We can observe a good match with the theory as d/R tends to 0.

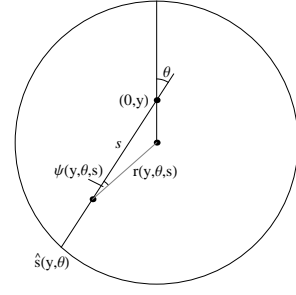


Fig. 9. Integration along a line in direction θ through the point $(0, y)$.

VI. SPATIAL DISTRIBUTION OF INFORMATION

Let us next consider the simplest possible mobility model where each user moves along a straight line (at least within the R -disk corresponding to the anchor zone) with a constant speed v . The user population is assumed homogeneous, with n nodes per unit area, and the velocity distribution isotropic, meaning that all directions of motion are equally probable.

Our aim is to find the criticality condition for this system. Furthermore, for systems which are above the threshold, the task is to find the function $p(r, \theta)$ defining which proportion of the nodes at distance r from the origin and moving in direction θ with respect to the *outward* radial direction are information-carrying “tagged” nodes.

A. Criticality Condition

Under the above assumptions, one can easily show that the mean number of encounters a test node, at distance r from the origin and moving in direction ψ with respect to the *inward* radial direction, experiences with tagged nodes per unit length of its path is

$$\tilde{\lambda}(r, \psi) = \frac{\lambda}{4} \int_0^{2\pi} p(r, \theta) \left| \cos \frac{\theta - \psi}{2} \right| d\theta, \quad (17)$$

where

$$\lambda = \frac{8}{\pi} n d \quad (18)$$

is the mean number of such encounters per unit length of path if all nodes were tagged, i.e., if $p(r, \theta) = 1$. The cosine factor in (17) reflects the fact that the encounter rate between two streams of nodes depends on their relative velocity, being zero for parallel streams and at maximum for opposing directions.

To determine the distribution $p(r, \theta)$ for the fluid system in stationarity consider the point $(0, y)$ at the distance y from the origin, see Fig. 9. Then we have

$$p(y, \theta) = 1 - \exp \left(- \int_0^{\hat{s}(y, \theta)} \tilde{\lambda}(r(y, \theta, s), \psi(y, \theta, s)) ds \right), \quad (19)$$

where

$$\begin{aligned} \hat{s}(y, \theta) &= \sqrt{R^2 - y^2 \sin^2 \theta} + y \cos \theta, \\ r(y, \theta, s) &= \sqrt{s^2 + y^2 - 2s y \cos \theta}, \\ \psi(y, \theta, s) &= \arccos \frac{s - y \cos \theta}{\sqrt{s^2 + y^2 - 2s y \cos \theta}}. \end{aligned} \quad (20)$$

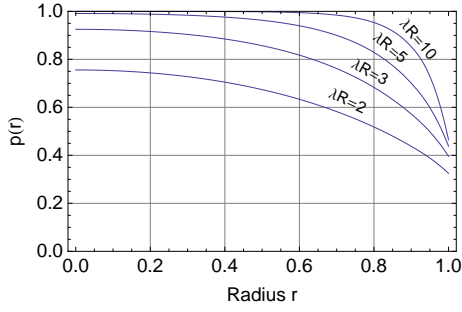


Fig. 10. Fixed point solutions for the distribution $p(r)$ as a function of $r = 0, \dots, 1$ in the cases (bottom up) $\lambda R = 2, 3, 5$, and 10 .

The integral in (19) is along the the line going through the point $(0, y)$ in direction θ . Variable s represents the distance of the point of integration from the point $(0, y)$. The exponential term is recognized to be the probability of no information exchange encounter between the point of entrance $\hat{s}(y, \theta)$ and the point of consideration $(0, y)$; its complement then gives the probability that there has been at least one such encounter, i.e., that the information has been gained. Expressions $r(y, \theta, s)$ and $\psi(y, \theta, s)$ give the distance of the point of integration from the origin and the angle that the line of integration makes with the inward radial vector at the point of integration (Fig. 9).

Since the function $\tilde{\lambda}(\cdot, \cdot)$ in the integrand depends on $p(r, \theta)$ by (17), equation (19) constitutes a non-linear integral equation for $p(r, \theta)$. This equation can be solved by iteration, by starting with an initial guess, e.g. $p(r, \theta) = 1$, substituting this on the right hand side of the equation thus obtaining a new function $p(r, \theta)$ (the variable name, r vs. y , of course, does not matter), and continuing until the fixed point solution is found.

Above the criticality point, the iteration converges quite fast, requiring only a few rounds. Solutions for $p(r) = \frac{1}{2\pi} \int_0^\pi p(r, \theta) d\theta$, i.e., the total proportion of tagged nodes of all the nodes at distance r , are shown in Fig. 10 for different values of λR . Corresponding full directional distributions $p(r, \theta)$ as a function of θ are shown in Fig. 11 for different values of r (curve parameter in each family of curves). From Fig. 10 we see that when λR becomes smaller, also the proportion $p(r)$ of tagged nodes decreases. It turns out that for a certain critical value of λR the solution goes to zero, which means that the information cannot be anymore sustained, even within a fluid model. The critical value can be most easily found by linearizing (19), justified when $p(r, \theta)$ is small, retaining only the first two terms of the Taylor series expansion of the exponential function,

$$p(y, \theta) = \int_0^{\hat{s}(y, \theta)} \tilde{\lambda}(r(y, \theta, s), \psi(y, \theta, s)) ds, \quad (21)$$

which, together with (17) constitutes an eigenvalue equation for λ . The critical value corresponds to the greatest eigenvalue, which can again be found by a standard iterative procedure, similar to that described above. In this way we find

$$\lambda_{cr} R = 1.038. \quad (22)$$

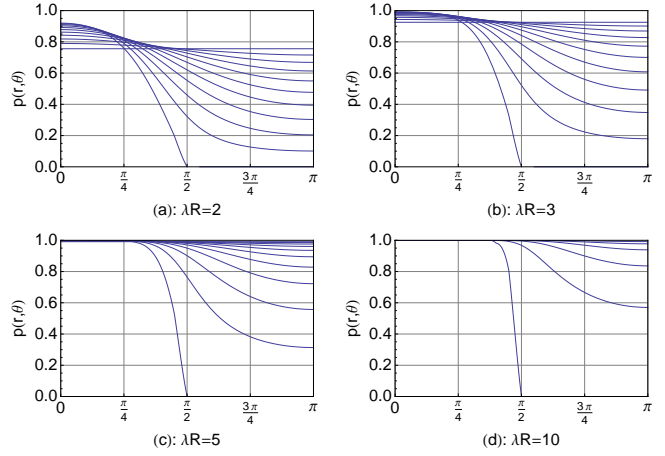


Fig. 11. Directional distribution of $p(r, \theta)$ at $r = 0.1, 0.2, \dots, 0.9$ (from the flattest curve to the steepest curve). Probability $p(r, \theta)$ is on the y -axis, and the meeting angle θ on the x -axis.

In terms of the more elementary parameters of (18), the criticality condition reads

$$n d R \geq 0.407. \quad (23)$$

For comparison, one can also analyse this mobility pattern similarly as in Section V. The mean sojourn time in the anchor zone is $1/\mu = \pi R/(2v)$, i.e., $\Lambda = 2R \cdot nv$. Moreover, $\Phi = nv$, $\varphi = nv/(2\pi)$, and (4) gives $\mathcal{R} = 4n^2 v d R^2 = \Lambda^2 d/v$. Omitting the spatial dependence and simply substituting these to the non-spatial criticality condition (8) gives

$$n d R \geq 1/4, \quad (24)$$

which is thus more optimistic estimate than (23).

1) *Example:* Simulation results with a transmission range of $d = 10$ m, the anchor zone radii of $R = 100$ m, 200 m, 400 m, and the node's velocity of 1 m/s, are illustrated in Fig. 12. The system is initialized by passing the information to the first two nodes arriving to the vicinity of the center, which prevents situations where a sole carrier walks immediately out of the anchor zone without passing the information further. From figure one can observe (i) that the criticality conditions, (23) in particular, correctly predict the region where the floating is feasible, and (ii) that the criticality condition becomes more accurate as the system dimensions approach the fluid limit.

B. Additional Results

From the previous model we can extract some more information. For notational simplicity we write the formulae for $R = 1$. The proportion of tagged nodes in the anchor zone is

$$p_{ave} = \frac{2}{\pi} \int_0^1 \int_0^\pi r p(r, \theta) d\theta dr. \quad (25)$$

If there are N nodes in total within the anchor zone, then the number of tagged nodes is $p_{ave} N$. Secondly, we can calculate the probability p_{inf} that a random node entering the anchor zone will acquire the information, i.e., exit as a tagged node,

$$p_{inf} = \int_0^{\pi/2} p(1, \theta) \cos \theta d\theta. \quad (26)$$

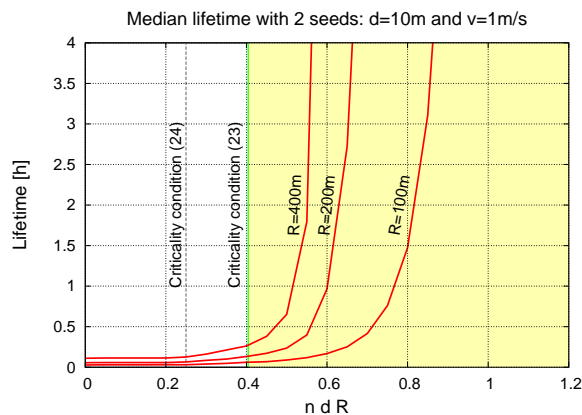


Fig. 12. Median information lifetime with uniform direct mobility model with three anchor zone radii: 100 m, 200 m and 400 m. Initially the information is passed to two nodes by a seeder located at the center. The two vertical lines correspond to criticality conditions (24) (left) and (23) (right).

Related to this, it is of interest to calculate the average length ℓ_{inf} traversed within the anchor zone upon information acquisition of a random node exiting as a tagged node,

$$\ell_{inf} = \frac{1}{p_{inf}} \int_0^1 \int_0^\pi \left(p\left(1, \arccos \sqrt{1-r^2 \sin^2 \theta}\right) - p(r, \theta) \right) r dr d\theta.$$

For comparison, one should note that the average length traversed within the anchor zone by a node that hits the anchor zone equals “the average height of the disk, i.e., the area of the disk π divided by the length of the diameter 2 or $\pi/2$.” Numerical results for these quantities are given in Table I.

VII. CONCLUSIONS

In this paper, we have studied the floating content concept, and in particular, the expected lifetime of the information in such a system. In the analysis, we have considered situations where the anchor zone is assumed to be significantly larger than the transmission range. When the diameter of the anchor zone is several magnitudes higher than the communication range between two nodes, the spatial dimension must be taken explicitly into account.

We develop a sound theoretical framework to analyze the floating content concept in this regime. Exact expressions for the fundamental quantities, such as the node encounter rate and mean contact times as a function of location are given. Our results allow one to characterize the feasible operation region ensuring sufficiently long expected information lifetimes. In particular, the so-called *criticality condition*, related to the mean number of encounters a randomly chosen node experiences per visit, provides a threshold above which the expected lifetime of the information increases very rapidly, and a content can actually “float” without any supporting infrastructure. The analytical results are applied to various mobility models, and supported by simulation experiments.

TABLE I
NUMERICAL RESULTS WITH SPATIAL MODEL.

λR	p_{ave}	p_{inf}	ℓ_{inf}/R
10	0.922	0.992	0.121
5	0.842	0.971	0.240
3	0.727	0.916	0.386
2	0.567	0.782	0.532
1.038	0	0	0.770

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] 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.
- [3] B. Liu, B. Khorashadi, D. Ghosal, C.-N. Chuah, and M. Zhang, “Assessing the VANET’s local information storage capability under different traffic mobility,” in *Proc. of IEEE INFOCOM*, 2010.
- [4] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. on Information Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.
- [5] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, “Closing the gap in the capacity of random wireless networks via percolation theory,” *IEEE Trans. on Information Theory*, vol. 53, no. 3, pp. 1009–1018, 2007.
- [6] R. Groenevelt, P. Nain, and G. Koole, “The message delay in mobile ad hoc networks,” *Performance Eval.*, vol. 62, no. 1-4, pp. 210–228, 2005.
- [7] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, “Performance modeling of epidemic routing,” *Computer Networks*, vol. 51, no. 10, 2007.
- [8] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnović, “Power law and exponential decay of inter contact times between mobile devices,” in *Proceedings of ACM MobiCom*, 2007, pp. 183–194.
- [9] Z. Kong and E. Yeh, “On the latency for information dissemination in mobile wireless networks,” in *Proc. of ACM MobiHoc*, 2008.
- [10] P. Jacquet, B. Mans, and G. Rodolakis, “Information propagation speed in mobile and delay tolerant networks,” *CoRR*, vol. abs/0903.1157, 2009.
- [11] E. Altman, T. Başar, and F. D. Pellegrini, “Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks,” in *ValueTools*, 2008.
- [12] E. Altman, G. Neglia, F. D. Pellegrini, and D. Miorandi, “Decentralized stochastic control of delay tolerant networks,” in *IEEE INFOCOM*, 2009.
- [13] K. M. Case, F. de Hoffmann, and G. Placzek, *Introduction to the theory of neutron diffusion. Volume 1*. Los Alamos, New Mexico: Los Alamos Scientific Laboratory, Apr. 1953.
- [14] S. Toumpis, “Mother nature knows best: A survey of recent results on wireless networks based on analogies with physics,” *Computer Networks*, vol. 52, no. 2, pp. 360–383, 2008.
- [15] E. Hyttiä and J. Virtamo, “On the optimality of field-line routing in massively dense wireless multi-hop networks,” *Performance Evaluation*, vol. 66, no. 3–5, pp. 158–172, Mar. 2009.
- [16] V. Lenders, M. May, G. Karlsson, and C. Wacha, “Wireless ad hoc podcasting,” *ACM/SIGMOBILE Mobile Computing and Communications Review*, Jan. 2008.
- [17] E. Hyttiä and J. Virtamo, “Random waypoint mobility model in cellular networks,” *Wireless Networks*, vol. 13, no. 2, Apr. 2007.
- [18] H. Cai, C.-H. Lee, and D. Y. Eun, “Invariance property of isotropic random walk mobility patterns in mobile ad-hoc networks,” in *Proc. of IEEE ICC*, 2008, pp. 2141–2145.
- [19] J. D. C. Little, “A proof of the queueing formula $L = \lambda W$,” *Operations Research*, vol. 9, no. 3, pp. 383–387, 1961.
- [20] P. Diaconis and L. Miclo, “On times to quasi-stationarity for birth and death processes,” *J. of Theor. Probability*, vol. 22, pp. 558–586, 2009.
- [21] 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.
- [22] 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.
- [23] J. L. Boudec, “On the stationary distribution of speed and location of random waypoint,” *IEEE Transactions on Mobile Computing*, Dec. 2005.
- [24] E. Hyttiä, J. Virtamo, P. Lassila, J. Kangasharju, and J. Ott, “Characterizing availability of anchored information in opportunistic content sharing,” Tech. Rep., Jul. 2010, <http://www.netlab.hut.fi/~esa/tr-06-10.pdf>.