# Upper Bounds on Expected Hitting Times in Mostly-Covered Delay-Tolerant Networks

Max F. Brugger, Kyle Bradford, Samina Ehsan, Bechir Hamdaoui, Yevgeniy Kovchegov
Oregon State University, Corvallis, OR 97331
bruggerm,bradfork,ehsans,hamdaoub,kovchegy@onid.orst.edu

***Abstract*— We derive theoretical bounds on expected hitting times in densely covered delay-tolerant networks (DTNs). We consider a number of fixed (data collector) nodes deployed in the DTN region, and a number of mobile (data generator) nodes that move freely in the region according to Brownian motion. As it moves, each mobile node is assumed to continuously generate and buffer data. When a mobile node comes within the communication coverage range of a data collector node, it downloads its buffered data to it. Otherwise, it keeps generating and buffering its data. In this paper, we derive analytic bounds on the amount of time a mobile node spends without communication coverage. Then, using these derived bounds, we derive sufficient conditions on node density that statistically guarantee that the expected hitting times remain below a given threshold.**

## I. Introduction

Due to their importance and wide range of applications, delay-tolerant networks (DTNs) have attracted considerable research focus, ranging from protocol design [1, 2] to connectivity and delay modeling and analysis [3–7]. The work in [4] uses continuum percolation theory [8] to show how delays in large wireless networks scale with the Euclidean distance between the sender and the receiver. Speed of information propagation has recently also been studied analytically for static [5, 9] as well as mobile [6, 7, 10] DTNs. The authors in [5] derived upper bounds on the maximum propagation speed in large-scale wireless networks, and those in [10] derived analytic upper bounds on information delay in large-scale DTNs with possible mobility and intermittent connectivity. Network connectivity has also been intensively studied, but mostly in the context of large-scale networks only. In [3], the authors derived an upper bound on the delay sufficient for disconnected networks to become connected through node mobility. The work in [4] derived the minimum node density required to ensure connectivity in large static networks.

In contrast, this work aims at deriving analytic upper bounds on the expected time a mobile node spends without communication coverage in mostly, but not fully, covered DTNs as a function of the communication coverage ratio; i.e, DTNs whose coverage ratio is close to one. Intermeeting times, defined as the time a mobile node spends before running into another node, have been derived in [11] for the generalized hybrid random walk mobility model. In this work, La [11] shows that the distribution of intermeeting times can be approximated by an exponential distribution when mobile nodes move independently from one another and when the probability of establishing communication links among nodes is relatively low. This result provides support for our use of a Brownian Motion model of a mobile node, which we will show also has approximately exponential intermeeting times under the assumptions of the Poisson Clumping Heuristic [12] (and described in more detail in Section III).

In our studied DTN model, a number of fixed nodes (also referred to as *access points*) are deployed in the DTN region, where mobile nodes (also referred to as *data generators*) move freely in the region by following a Brownian motion. As it moves, each mobile node is assumed to continuously generate and buffer data. When a mobile node comes within the communication coverage range of a data collector node, it downloads all of its buffered data to it. In this work, we first use the Poisson Clumping Heuristic [12] to provide analytic bounds on the expected hitting time, the time a mobile node spends without communication coverage. Then, using these derived bounds, we derive sufficient conditions on node density that ensure that the expected hitting times are guaranteed to be below a given time threshold. Finally, using simulations, we validate/verify the derived results. Precisely, our contributions in this paper are: we

- Derive analytic bounds on the expected time mobile nodes spend without communication coverage.
- Provide sufficient node density conditions ensuring that the expected time mobile nodes spend without coverage remains below a fixed threshold.
- Validate/verify the derived results via simulations.

The rest of the paper is organized as follows. In Section II, we state our network model. To introduce our methods, we first overview the Poisson Clumping Heuristic in Section III. We then present our analytic results in Section IV. In Section V, we validate via simulations the derived bounds. Finally, we conclude in Section VI.

## II. Model

A number of fixed nodes (data collectors) are deployed in the DTN region, where mobile nodes (data generators) move freely in the region, following a Brownian motion. As they move, mobile nodes are assumed to continuously generate and buffer data independently from one another. When a mobile node comes within the communication coverage range of a data collector node, it downloads all of its buffered data to it. Otherwise, it keeps generating and buffering its data until it goes by a collector node.

Our focus in this work is on the study of dense DTNs. That is, DTNs that are mostly covered, but not fully. Hence, the network formed by the data collector nodes is assumed to be unconnected, and the communication coverage ratio is assumed to be close to 1. In these dense DTNs, as mobile nodes move, they will eventually traverse a data collector's communication coverage area, and can then download their buffered data. To this end, the *coverage ratio*[1] is assumed to be close to 1 throughout this paper, and all the mathematical analysis in this work depends heavily on this assumption.

## III. Poisson Clumping Heuristic

Given a time-dependent stochastic process and a set $A$, if the process intersects the set $A$ rarely, then we can approximate the behavior of this process' arrivals to the set by the Poisson Process. The Poisson Process states that the inter-arrival times of the mobile node are exponentially distributed, with parameter $\lambda$, and that the number of times the mobile node has hit an access point up to time $t$ is Poisson distributed, with parameter $\lambda t$. In the language of the heuristic, $\lambda$ is called the clump rate, so named because the random sets of times that the process spends in $A$, denoted $C$, appear to "clump" together. The approximations given by the Poisson Clumping Heuristic improve if the process is unlikely to return to $A$ immediately after leaving $A$; there should typically be some drift away from $A$. Let $\pi$ be the probability (of the stationary distribution) that the process is in $A$. Then, the main result of the Poisson Clumping heuristic is: $\pi = \lambda \mathbb{E}C$, where $\mathbb{E}C$ the expected size of $C$. The assumption that the interarrival times follow an exponential distribution additionally gives us that $\lambda = 1/T$, where $T$ is the hitting time. A detailed description/explanation of the theory behind the Poisson Clumping Heuristic can be found in [12].

## IV. Expected Hitting Time

We consider that mobile nodes follow a Brownian motion and move in a plane. Collector nodes are placed in the plane to form a grid. We assume that each of the collector nodes has a circular coverage region with radius $\kappa$. The spacing distance $D$ between two neighboring collector nodes is assumed to be larger than $\sqrt{2}\kappa$. This distance is also assumed to be smaller than $2\kappa$ so as to ensure that the DTN contains regions of no coverage, referred to as uncovered regions, that are disconnected. As shown in Fig. 1, we then draw a square of side length $D$ around each uncovered region, with the center of each region placed in the middle of the square. Each corner of each square corresponds to one collector node.

We make two observations before proceeding with our derivation and analysis. First, the problem is symmetric,

[1]The coverage ratio is defined as the fraction of the area covered by collector nodes' communication ranges to that of the total DTN area.

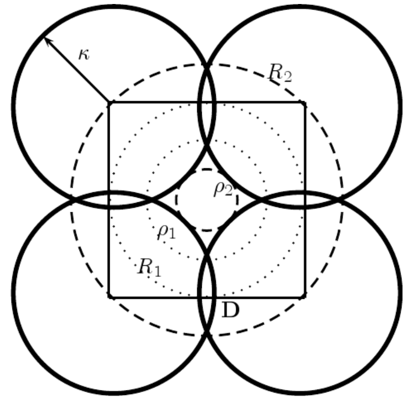


Fig. 1. Grid geometry: the uncovered area (the star shape) is bounded by two circles, one of radius $\rho_2$ from within and one of radius $\rho_1$ from without, which will be used to calculate bounds on the expected time a mobile node spends outside the communication coverage area. $D$ is the distance between two neighboring collector nodes; $\kappa$ is the radius of the communication coverage area of a collector node.

and hence studying one square suffices. Note that once a mobile node reaches the edge of a square coming from the edge of an uncovered region, returning back to the same region or another one makes no difference vis-a-vis of our clumping analysis. Second, because the uncovered region has an odd shape and our boundary region has a square shape, it is too difficult to derive the exact clump rate. Instead, we derive bounds on the clump rate.

We inscribe the largest possible circle in the uncovered region, centered at the center of the square, and denote the radius of the circle $\rho_1$. We also circumscribe the smallest possible circle around the uncovered region, centered at the center of each region, and denote the radius of this circle $\rho_2$. The geometry of this grid is shown in Fig. 1.

In our model, the distribution of hitting times satisfies the assumptions stated in Section III regarding the rarity with which the mobile node hits the uncovered area, because of 1) the assumption of a high coverage ratio, and 2) the fact that the drift for the radial part of Brownian Motion, given by the Bessel Process, has drift $\mu(r) = 1/(2r)$, where $r$ is the Euclidean distance of the Brownian Motion from the origin.

Let $\pi(C)$ be the probability (given by the stationary distribution) that the process is in the uncovered region and $\mathbb{E}C$ the expected amount of time spent in the uncovered region. As described in Section III, it follows that $\pi(C) = \lambda \mathbb{E}C$.

The time a mobile node spends with communication coverage corresponds to the time it takes a mobile node to reach the edge of the square from the edge of an uncovered region, and then to return to one of the uncovered regions again. To derive an upper bound, we first investigate a radial diffusive process on an inner disk with radius $\rho_1 = \frac{1}{2}D - \sqrt{\kappa^2 - \frac{1}{4}D^2}$ and an outer disk with radius $R_1 = \frac{1}{2}D$ centered at the same point (both circles are shown in Fig. 1). We then investigate a radial diffusive process on an inner disk with radius $\rho_2 = \frac{\sqrt{2}}{2}D - \kappa$ and an outer disk with radius $R_2 = \frac{\sqrt{2}}{2}D$ to find a lower bound on the clump rate. We can see that with these two disks, we are inscribing the square within a disk and another within the square, as is shown in Fig. 1. The geometric properties of the square allow for a probabilistic coupling construction, where the radial Brownian motion on a square region is stochastically dominated from above and below by the radial Brownian motions (Bessel processes) whose boundary conditions are respectively the inner and outer circle.

Since we are investigating the radial diffusive process on a disk with radius $R_i$ centered at the center of a circle of radius $\rho_i$ where $i \in \{1, 2\}$, the Brownian motion in either case can be modeled as a Bessel process with parameter 2, i.e., with drift $\mu(r) = \frac{1}{2r}$ and variance $\sigma^2 = 1$ [13].

### A. Inner and Outer Disks of Radii $\rho_i$ and $R_i > \rho_i$

Let us now consider a disk of radius $\rho_i$ centered in a disk of radius $R_i > \rho_i$ as shown in Fig. 1, where $i \in \{1, 2\}$. Let us assume that a mobile node moves inside the disk of radius $R_i$, and it bounces back when it hits the boundary of the disk. We define then the smaller disk of radius $\rho_i$ to be our clump, and the hitting time to be the time between two consecutive clump visits. The hitting time is then the time it takes a mobile node to reach the boundary of the outer disk of radius $R_i$ given it just left the inner disk of radius $\rho_i$ plus the time it takes a mobile node to hit back the inner disk given that it just bounced back from hitting the boundary. The expected value of this hitting time is stated in the following proposition (refer to Appendix B of [14] for proof).

*Proposition 4.1:* For a smaller disk of radius $\rho_i$ centered in the larger disk of radius $R_i$, the expected hitting time is $h(\rho_i, R_i)$, where $h(\rho, R) = R^2 \ln \left|\frac{R}{\rho}\right|$.

The following corollary, which is an immediate application of the hitting time function, $h$, is used throughout to derive the main results of this paper.

*Corollary 4.2:* For a star-shaped inner region and a square boundary of side length $D$, the expected hitting time is lower bounded by $h(\rho_1, R_1)$ and upper bounded by $h(\rho_2, R_2)$.

### B. Upper Bounds and Sufficient Conditions

We now derive upper bounds on the expected hitting times, and provide sufficient conditions on node density that guarantee that the expected time the mobile node spends without coverage does not exceed a given threshold. We define the node density $\nu$ to be equal to $1/D^2$.

*Proposition 4.3:* The expected amount of time a mobile node spends without communication coverage, $\mathbb{E}C$, is bounded above by

$$\frac{\kappa}{\sqrt{2\nu}}\left(1 - 2\sqrt{\kappa^2\nu - \frac{1}{4}}\right).$$

*Proof:* The probability of being in the uncovered region, $\pi(C)$, is the ratio of the expected amount of time the mobile node spends in the clump, $\mathbb{E}C$, to the expected amount of time the mobile node spends between clumps, $\mathbb{E}T$. More formally, $\pi(C) \approx \frac{\mathbb{E}C}{\mathbb{E}T}$. From Corollary 4.2, it follows then that

$$\pi(C) \cdot h(\rho_1, R_1) \leq \mathbb{E}(C) \leq \pi(C) \cdot h(\rho_2, R_2).$$

From a geometric argument, we can find the area of the uncovered region and divide it by the area of the square surrounding it to find $\pi(C)$,

$$\pi(C) = 1 - 2\sqrt{\kappa^2\nu - \frac{1}{4}} - \pi\kappa^2\nu + 4\cos^{-1}\left(\frac{1}{2\sqrt{\kappa^2\nu}}\right)\cdot\kappa^2\nu,$$

where the $\pi$ in the right hand side of the equation is the constant and the $\pi$ on the left hand side of the equation is the terminology for the stationary distribution.

Because $\frac{1}{4} \leq \kappa^2\nu \leq \frac{1}{2}$, we can then write

$$\pi(C) \leq 1 - 2\sqrt{\kappa^2\nu - \frac{1}{4}}.$$

Now since

$$-\frac{1}{2\nu}\ln\left(1 - \sqrt{2\kappa^2\nu}\right) \leq \frac{\kappa}{\sqrt{2\nu}},$$

then,

$$\mathbb{E}(C) \leq \frac{\kappa}{\sqrt{2\nu}}\left(1 - 2\sqrt{\kappa^2\nu - \frac{1}{4}}\right).$$

■

*Corollary 4.4:* For a sufficiently small threshold, $\tau$, (where we require $\tau \leq 2\kappa^2$ for the square root to remain real), the expected time a mobile node spends without communication coverage is guaranteed to remain below the threshold (i.e., $\mathbb{E}C \leq \tau$) if the density $\nu$ of collector nodes satisfies

$$\nu \geq 2 \cdot \left(\frac{\kappa}{\tau + \sqrt{4\kappa^4 - \tau^2}}\right)^2.$$

*Proof:* Proposition 4.3 provides an upper bound on $\mathbb{E}C$, so it suffices that

$$\frac{\kappa}{\sqrt{2\nu}}\left(1 - \sqrt{4\kappa^2\nu - 1}\right) \leq \tau$$

in order for the expected time to remain below $\tau$.

By letting $\hat{\nu} = \kappa^2\nu$, the above inequality becomes

$$\frac{\kappa^2}{\sqrt{2\hat{\nu}}}\left(1 - 2\sqrt{\hat{\nu} - \frac{1}{4}}\right) \leq \tau.$$

Now, solving for $\hat{\nu}$, we find

$$\sqrt{\hat{\nu} - \frac{1}{4}} \geq -\frac{1}{2}\left(\frac{\sqrt{2\hat{\nu}}\tau}{\kappa^2} - 1\right).$$

Because $\frac{1}{4} \leq \hat{\nu} \leq \frac{1}{2}$ and $\tau \leq \kappa^2$, we have that

$$-\frac{1}{2}\left(\frac{\sqrt{2\hat{\nu}}\tau}{\kappa^2} - 1\right) \geq 0.$$

This implies that

$$\left(\frac{\tau^2}{2\kappa^4}-1\right)\left(\sqrt{\hat{\nu}}\right)^2-\frac{\sqrt{2}\tau}{2\kappa^2}\sqrt{\hat{\nu}}+\frac{1}{2}\leq 0.$$

The leading coefficient is negative, so this is a parabola in $\sqrt{\hat{\nu}}$ that opens downwards. Using the quadratic formula to find the roots of this polynomial in $\sqrt{\hat{\nu}}$, (note one root will be negative, which is impossible for a square root, so we only need to concern ourselves with the positive root),

$$\sqrt{\hat{\nu}}\geq\frac{\frac{\sqrt{2}\tau}{2\kappa^2}-\sqrt{\frac{\tau^2}{2\kappa^4}+2\left(1-\frac{\tau^2}{2\kappa^4}\right)}}{2\left(\frac{\tau^2}{2\kappa^4}-1\right)}$$

which implies that

$$\hat{\nu}\geq 2\cdot\left(\frac{1}{\frac{\tau}{\kappa^2}+2\sqrt{1-\frac{1}{4}\left(\frac{\tau}{\kappa^2}\right)^2}}\right)^2.$$

Replacing $\hat{\nu}$ by $\nu\kappa^2$ results in the sufficient node density stated in the corollary, which guarantees that the expected time a mobile node spends without communication coverage will be less than the threshold, $\tau$. ■

## V. Simulations

In this section, we first validate the use of the Poisson Clumping Heuristic by measuring the expected hitting times and comparing them against the derived bounds, and then verify the derived sufficient conditions on node density by mimicking and simulating a Brownian motion.

### A. *Poisson Clumping Heuristic Validation*

Recall that, as illustrated in Section III, the derived theoretical results are based on the assumption that Brownian motion in dense networks yields approximately exponentially distributed intermeeting times (i.e., times without communication coverage are exponentially distributed), thus allowing us to use the Poisson Clumping Heuristic approach. In this section, we focus on validating the Poisson Clumping Heuristic approach by simulating and measuring the hitting times of a Brownian motion, and comparing them with the theoretical upper bound (in Corollary 4.2) for a fixed value of $\kappa$. Recall that $D$ can range, for fixed $\kappa$, from $2\kappa$, where the circles of the covered regions just barely touch, to $\sqrt{2}\kappa$, where the circles overlap so that the uncovered region is at its smallest; the coverage ratio $\eta$ varies respectively from about 0.7854 to 1. In our simulation, we set $\kappa = 5$.

We use Matlab to simulate a Brownian motion in a square by generating a normal random variable with distribution $(\mu = 0, \sigma^2 = 1)$ for the displacement of the mobile node, and a uniform random number selected from $[0, \pi]$ for the angle the mobile node's path makes with the x-axis.

Because of symmetry, simulating a Brownian motion on a plane is equivalent to simulating it on a square of side length $D$ with 4 collector nodes each located at one corner and an uncovered area located at its center. The

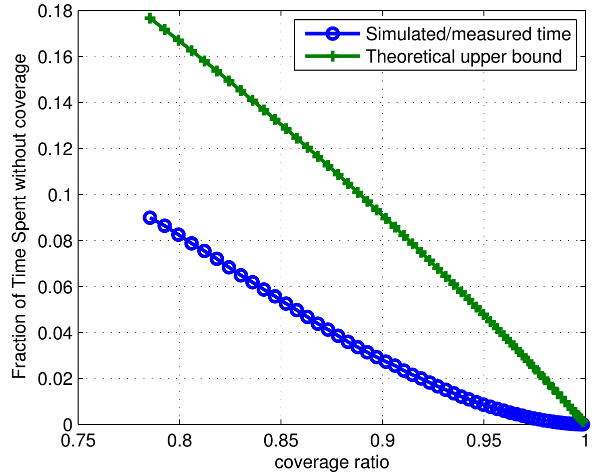


Fig. 2. Measured and theoretical hitting times for a Brownian motion on a square with $\kappa = 5$ and $D$ ranging from $\sqrt{2}\kappa$ to $2\kappa$.

covered region then is the area within $\kappa$ of any corner. If the simulated Brownian motion exits the square, it is equivalent to continue the simulation with the mobile node placed back inside the box at the opposite position. Fig. 2 shows that the simulated times are well bounded by the derived upper bounds for a range of values for $\eta$, and as expected, the higher the coverage, the tighter the bound.

### B. *Sufficient Node Density Conditions Verification*

In this section we test and verify the sufficient conditions on node density stated in Proposition 4.3 to ensure that the expected time a mobile node spends without communication coverage is guaranteed to remain below the threshold.

We discuss here how to determine which values of the thresholds are appropriate for consideration with our heuristic. First, we fix values for $\kappa$ and $D$. Because the expected amount of time the mobile node spends in the uncovered region is related to the size of the uncovered region relative to the side length, $D$, of the square, and thus the hitting time, and because the Poisson Clumping Heuristic most closely approximates the behavior of our process when these hitting times are large, we choose

$$\tau = \pi(C)/0.08 \cdot T$$

where $T$ is in the 99th percentile of an exponential distribution function with mean $h(\rho_2, R_2)$, which is the distribution the heuristic assumes the hitting times fall into. The $1/0.08$ term accounts for the step size in the simulation.

We verify the sufficient condition for the expected time a mobile node spends without communication coverage to be guaranteed to remain below a threshold, $\tau$, for a range of values of $\kappa$: 2, 5, 10 and 20, and for $D$ near to $\sqrt{2}\kappa$ for each ($1.45\kappa$ to $1.78\kappa$). We calculate the sufficient density from Proposition 4.3, which is a function of $\tau$ and $\kappa$. For each ratio of density to sufficient density (which is varied by varying the values of $D$), we simulate and measure the time

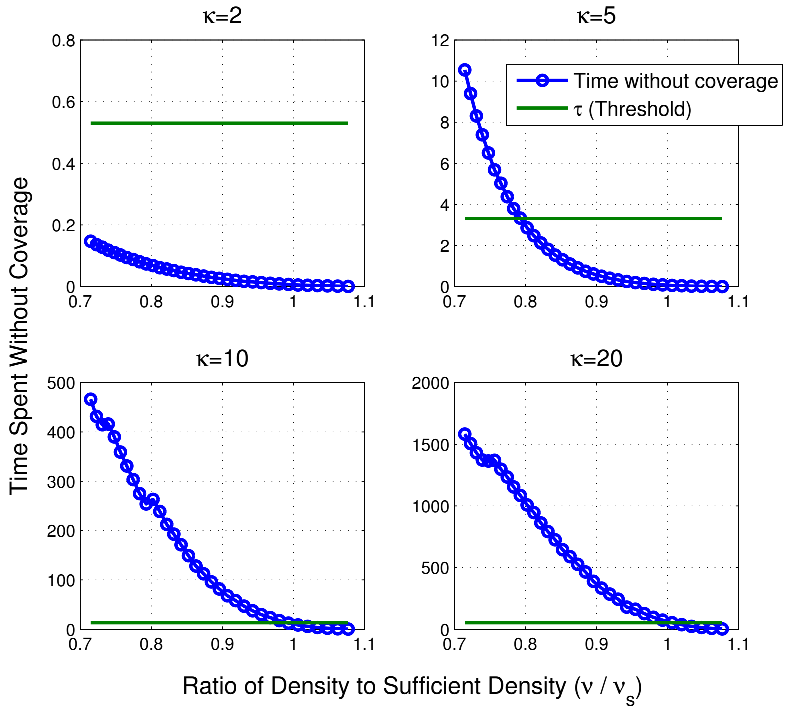


Fig. 3. The measured and the derived upper bound (Proposition 4.3) on the time a mobile node spends without communication coverage when varying $D$ at values near $\sqrt{2}\kappa$, for $\kappa = (2, 5, 10, 20)$.

spent outside the coverage for each value of $\kappa$. The results are presented in Fig. 3. Observe that when the ratio of density to sufficient density ($\nu/\nu_s$) is greater than one (i.e., the sufficient density condition is met), the time a mobile node spends without communication coverage is below the threshold $\tau$.

To summarize, through simulations, we are able to first validate/support the use of the Poisson Clumping Heuristic techniques and then illustrate/verify the applicability of the derived sufficient conditions on the node density.

## VI. Conclusion

In this paper, we derived theoretical bounds on expected hitting times in mostly covered DTNs. We first provided analytic bounds/approximations on the expected time mobile nodes spend without communication coverage, and then derived sufficient conditions ensuring that these times are guaranteed to remain below a given threshold. Finally, we verified our derived results via simulations.

## References

[1] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *Proc. of INFOCOM*, 2010.

[2] T. Spyropoulos, T. Turletti, and K. Obraczka, "Routing in delay-tolerant networks comprising heterogeneous node populations," *IEEE Trans. on Mobile Computing*, August 2009.

[3] F. De Pellegrini, D. Miorandi, I. Carreras, and I. Chlamtac, "A graph-based model for disconnected ad hoc networks," in *Proc. of INFOCOM*, 2007.

[4] Z. Kong and E. M. Yeh, "Connectivity and latency in large-scale wireless networks with unreliable links," in *Proc. of INFOCOM*, 2008.

[5] Y. Xu and W. Wang, "The speed of information propagation in large wireless networks," in *Proc. of INFOCOM*, 2008.

[6] Z. Kong and E. Yeh, "On the latency for information dissemination in mobile wireless networks," in *Proc. of ACM MOBIHOC*, 2008.

[7] P. Jacquet, B. Mans, and G. Rodolakis, "On space-time capacity limits in mobile and delay tolerant networks," in *Proc. of INFOCOM*, 2010.

[8] R. Meester and R. Roy, "Continuum percolation," *New York: Cambridge University Press*, 1996.

[9] R. Zheng, "Information dissemination in power-constrained wireless networks," in *Proc. of INFOCOM*, 2006.

[10] P. Jacquet, B. Mans, and G. Rodolakis, "Information propagation speed in mobile and delay tolerant networks," in *Proc. of INFOCOM*, 2009.

[11] R. J. La, "Distributed convergence of intermeeting times under the generalized hybrid random walk mobility model," *IEEE Transactions on Mobile Computing*, September 2010.

[12] D.J. Aldous, *Probability approximations via the Poisson clumping heuristic*, Springer-Verlag New York, 1988.

[13] S. Karlin and H.M. Taylor, *A second course in stochastic processes*, Academic press, 1999.

[14] K. Bradford, M. Brugger, S. Ehsan, B. Hamdaoui, and Y. Kovchegov, "Data loss modelling and analysis in partially-covered delay-tolerant networks," in *Proc. of 20th International Conference on Computer Communications and Networks (ICCCN)*, 2011.