The Node Distribution of the Random Waypoint Mobility Model for Wireless Ad Hoc Networks*

Christian Bettstetter† Giovanni Resta‡ Paolo Santi‡

Abstract

The random waypoint model is a commonly used mobility model in the simulation of ad hoc networks. It is known that the spatial distribution of network nodes moving according to this model is in general non-uniform. However, a closed-form expression of this distribution and an in-depth investigation is still missing. This fact impairs the accuracy of the current simulation methodology of ad hoc networks and makes it impossible to relate simulation–based performance results to corresponding analytical results.

To overcome these problems, we present a detailed analytical study of the spatial node distribution generated by random waypoint mobility. More specifically, we consider a generalization of the model, in which the pause time of the mobile nodes is chosen arbitrarily in each waypoint and a fraction of nodes may remain static for the entire simulation time. We show that the structure of the resulting distribution is the weighted sum of three independent components: the static, pause, and mobility component. This division enables us to understand how the model's parameters influence the distribution. We derive an exact equation of the

*Manuscript submitted 6 Sep. 2002; revised 28 Apr. and 18 Jul. 2003.

†Technische Universität München, Inst. of Communication Networks (LKN), D-80290 Munich, Germany. Ph: +49-89-289-25813. Fax: -23523. Email: `bettstetter@ei.tum.de`. Web: `http://www.lkn.ei.tum.de/~chris`

‡Istituto di Informatica e Telematica, CNR, Area della Ricerca di San Cataldo, Via G. Moruzzi 1, 56124 Pisa, Italy. Ph: +39-050-3152-411. Fax: -333. Email: {`giovanni.resta, paolo.santi`}`@iit.cnr.it`

asymptotically stationary distribution for movement on a line segment and an accurate approximation for a square area. The good quality of this approximation is validated through simulations using various settings of the mobility parameters. In summary, this article gives a fundamental understanding of the behavior of the random waypoint model.

Index Terms: Mobility modeling, random waypoint model, mobile ad hoc networking, simulation

1 Introduction

Performance analysis in presence of mobility is of major importance in the design of wireless communication and computer networks. Since real movement patterns are difficult to obtain, a common approach is to use synthetic mobility models, which resemble to some extent the behavior of real "mobile entities" (see [2, 10, 13, 15, 19, 22, 36]). Based on such models, basic conclusions with respect to critical network parameters can be provided.

The most commonly used mobility model in the ad hoc networking research community is the *random waypoint (RWP) model* [21]. It is implemented in the simulation tools NS2 [25] and GloMoSim [35] and used in many evaluations of network algorithms and protocols (see [9, 11, 18]). In this stochastic model, each node of the network chooses uniformly at random a destination point ("waypoint") in a rectangular deployment region Q. A node moves to this destination with a velocity v chosen uniformly at random in the interval $[v_{min}, v_{max}]$. When it reaches the destination, it remains static for a predefined pause time t_p, and then starts moving again according to the same rule.

It has been observed in [2, 6–8] and [28] that the spatial distribution of nodes moving according to the RWP model is non-uniform. Although the initial node positioning is typically taken from a uniform random distribution, the mobility model changes this distribution during the simulation. This effect, known as border effect [2], occurs because nodes tend to cross the center of Q with a relatively high frequency. For a long running time of the movement process, the stochastic distribution of the nodes converges toward an asymptotically stationary distribution with the maximum node density in the middle of Q.

The non-uniformity of the RWP node distribution has important practical consequences. First, it reduces the applicability of existing analytical results concerning ad hoc networks, which are typically based on the uniformity assumption. For example, theoretical results with respect to routing ([1, 24]), capacity ([14, 17]), connectivity ([3, 16, 30]), and minimum power issues cannot be applied directly in a mobile scenario that employs the RWP model. Second, the non-uniform distribution implies that the representativeness of the huge amount of simulation results obtained by using the RWP model could be impaired. This is because the short-term behavior of the RWP model is quite different from the actual long-term behavior. To overcome these problems, this article investigates in detail the RWP node distribution as a function of the mobility parameters.

In fact, we consider a generalized version of the RWP model. In this generalized model, a node may remain static for the entire simulation time with a given probability. Hence, only a fraction of the nodes are expected to move. Furthermore, we consider the fact that nodes are initially distributed according to an arbitrary spatial distribution. Last, but not least, we allow the pause time t_p to be different after each movement period.

The rest of this article is organized as follows. Section 2 outlines related work and motivates in more detail our interest in the derivation of the RWP node distribution. Section 3 motivates and explains the introduction of the generalized RWP model as described above. We formally characterize this model as a stochastic process and discuss some of its properties that are useful in the derivation of the node distribution. Furthermore, we show that this distribution is the sum of three distinct components: the static, pause, and mobility component. This separation enables us to understand the influence of the model's parameters on the resulting long-term node distribution. Next, in Section 4, we study in detail the mobility component of the distribution, i.e., the component that results when all nodes are continuously moving ($t_p = 0$, no static nodes). We derive an exact equation for RWP movement on a line segment and an accurate approximation for movement on a square area. In Section 5, we characterize the static and pause components and present the expression of the overall node distribution. In Section 6, several simulation results show that the approximation used in the derivation of the mobility component on a square is negligible in practice. Finally, Section 7 summarizes our contributions.

2 Motivation and Related Work

Despite the popularity of the RWP model, an in-depth understanding of its behavior is still lacking in the community. Only recently some papers appeared that study its stochastic properties and warn researchers of pitfalls that might occur when using this model (see [2, 5, 7, 8, 10, 28, 31, 34]). Probably the first simulation–based studies of the spatial node distribution were made in [2] and [8]. The fact that the long-term node distribution is different from the initial uniform distribution, calls into question the representativeness of many simulation results in the literature. Typical settings for a simulation-based analysis of ad hoc networks are the following: a few tenths or a hundred of nodes are distributed uniformly at random in a rectangular region; then, they start moving according to the RWP model. The behavior of the mobile network is observed for a number of time steps (where one step often corresponds to one second) in the order of, at most, one thousand. Such settings have been used, for instance, in the evaluation of routing ([11, 12, 18, 20, 21, 32]), multicast [27], and energy-conserving [33] protocols. Given the typical values of the mobility parameters used in the simulations, it follows that nodes in the above described scenario perform in general only a very limited number of movement periods during the simulation time. These are in general not enough to reach the "steady state" of the network. In other words, *observing the network for relatively few steps after the initial node positioning is not representative for the actual long-term behavior of the system.*

The lack of accuracy of the methodology which is currently used to simulate ad hoc networks has also been outlined in a recent paper [34] from a different perspective. The authors show that the average of the nodes' speed decreases over time and converges to a value $\bar{v}$ that is strictly less than the initial average speed $\frac{v_{min}+v_{max}}{2}$ (unless $v_{min} = v_{max} = v > 0$). Furthermore, setting $v_{min} = 0$ (as it is done in many simulations of ad hoc networks [11, 12, 20, 27, 33]) is particularly critical, since in this case $\bar{v}$ is arbitrarily close to zero, and the mobile system will eventually converge to an almost static one. The authors perform several experiments to support their argumentation, showing that the performance of commonly used routing algorithms can vary considerably with time: typically, after an initialization phase, whose duration depends on the values of v_{min} and

v_{max}, the performance of the protocol converges toward the "steady-state performance."

By giving the steady-state distribution of RWP nodes, this article is a further step in the direction of improving the accuracy of ad hoc network simulations.

3 Definition of Generalized RWP Movement

3.1 Parameters

The following parameters describe a simulation setup with generalized RWP mobility in a complete manner:

– Size and shape of the deployment region Q

– Initial spatial node distribution $f_{init}(\mathbf{x})$

– Static parameter p_s, with $0 \leq p_s \leq 1$

– Probability density function $f_{T_p}(t_p)$ of the pause time

– Minimum speed and maximum speed: $0 < v_{min} \leq v_{max}$

In this article, we consider one and two-dimensional deployment regions of the form $Q = [0, a]^\alpha$ with $\alpha = 1, 2$. The initial node distribution $f_{init}(\mathbf{x})$ is used to place nodes at the beginning of a simulation in Q. In general, it is different from a uniform distribution.

The parameter p_s represents the probability that a node remains static for the entire simulation time. This accounts for all situations in which a fraction of the nodes are not able to move. This could be the case if sensors are spread from a moving vehicle, and some of them remain entangled, say, in a bush or tree. This can also model a situation in which two types of nodes are used: one type is static, and another type is mobile. To a certain extent, using a separate parameter to model static nodes solves the pitfall described in Section 2 that arises when v_{min} is set to 0 as done in many papers. The rationale for setting $v_{min} = 0$ was to allow some of the nodes to be "almost static." Unfortunately, this implies that *all* the network nodes will eventually become almost static [34], which seems to be quite unrealistic in many application scenarios. In our extended RWP model, we thus explicitly separate the static and the mobile part of the network.

3.2 Stochastic Movement Process

Considering the mobile part of the network, we note that each RWP node moves independently of other nodes. Thus, all nodes have the same stochastic movement properties, and we can concentrate our attention on a single RWP node: its asymptotic spatial distribution is the same as the asymptotic distribution of all nodes.

The movement periods of a node are indexed by the discrete-time parameter i, where $i \in \mathbb{N}$, and the continuous time is denoted by t. The following random variables are used:

- Destination point $\mathbf{D}_i$ in its α-dimensional coordinates
- Pause time $T_{p,i}$ in the destination point $\mathbf{D}_i$
- Velocity V_i of the node during period i

With these definitions, the RWP model can be formally described as a stochastic process

$$\{\mathbf{D}_i, T_{p,i}, V_i\}_{i\in\mathbb{N}} = \{(\mathbf{D}_1, T_{p,1}, V_1), (\mathbf{D}_2, T_{p,2}, V_2), \ldots\} ,$$

where an additional waypoint $\mathbf{D}_0$ is needed for initialization. A sample of the process is denoted by $\{\mathbf{d}_i, t_{p,i}, v_i\}_{i\in\mathbb{N}}$. One movement period i is completely defined by the set $\{(\mathbf{d}_{i-1}, \mathbf{d}_i), t_{p,i}, v_i\}$.

We always assume that the random waypoints $\mathbf{D}_i$ are independently and identically distributed (i.i.d.) at random using the uniform distribution over Q. Only the initial waypoint $\mathbf{D}_0$ is determined by $f_{init}(\mathbf{x})$. The movement vector from $\mathbf{d}_{i-1}$ to $\mathbf{d}_i$ is denoted as *trajectory* τ_i. The complete movement trace of a node can thus be described by the sequence of these trajectories, i.e.,

$$\begin{aligned} &\{\tau_1, \tau_2, \ldots, \tau_i, \ldots\} = \\ &\quad = \{\mathbf{d}_0 - \mathbf{d}_1, \mathbf{d}_1 - \mathbf{d}_2, \ldots, \mathbf{d}_{i-1} - \mathbf{d}_i, \ldots\} . \end{aligned}$$

As an alternative to the random variable $\mathbf{D}_i$, we also use the random variable $\mathbf{S}_i$ denoting the starting waypoint of the i-th movement period. Clearly, the starting point of the current period is the destination point of the previous period, i.e., $\mathbf{S}_i = \mathbf{D}_{i-1}$. Again, we use lower case notation for samples of the process. When we just refer to a single random variable of the process, we omit the index i.

Regarding the pause times $t_{p,i}$, the original RWP model forces the nodes to have the same pause time t_p in each waypoint during the entire movement process, i.e., $T_{p,i} = t_p = const. \forall i$. This is a rather unrealistic aspect of RWP mobility, which is further amplified by the fact that the pause time is assumed to be the same for all the nodes in the network. In our generalized model, we assume that the pause time after each movement period is chosen from an arbitrary pdf $f_{T_p}(t_p)$ in the interval $[t_{p,min}, t_{p,max}]$ with $t_{p,min} \geq 0$ and a well-defined expected value $E[T_p]$. This distribution does not change over time and is the same for all the nodes in the network. Observe that our probabilistic homogeneity assumption is far less stringent than the equality assumption of the original model.

In each waypoint a node chooses a new speed V_i uniformly at random from the interval $[v_{min}, v_{max}]$. We explicitly request that $v_{min} > 0$ to avoid deadlocks in the movement process.

3.3 Ergodicity Properties

In the derivation of the spatial distribution, the distances between two consecutive waypoints, i.e., the trajectory lengths $l_i = \|\tau_i\| = \|\mathbf{d}_i - \mathbf{d}_{i-1}\|$, play an essential role. While the random waypoints are independent by definition, these random lengths are not stochastically independent; in fact, the endpoint of one movement period is the starting point of the next movement period. Instead of considering a chained set of trajectories, we consider a set of independent and disjoint trajectories between pairs of independent random points, i.e.,

$$\begin{aligned} &\{\tau'_1, \tau'_2, \ldots, \tau'_i, \ldots\} = \\ &\qquad = \{\mathbf{d}'_0 - \mathbf{d}'_1, \mathbf{d}'_2 - \mathbf{d}'_3, \ldots, \mathbf{d}'_{2i-12} - \mathbf{d}'_{2i-1}, \ldots\}\,, \end{aligned}$$

where the points are uniformly distributed in Q. We claim that several statistical properties are shared by this independent random point (IRP) process and the RWP process.

Let us consider a function $z(\tau)$ of the two endpoints of a trajectory (e.g., the trajectory length $z(\tau) = \|\tau\|$), and let us denote the corresponding random variables in the RWP and IRP process by Z and Z', respectively. We want to show that $E[Z] = E[Z']$. To do so, we consider an infinite RWP trace $\{\tau_1, \tau_2, \ldots\}$ and an infinite IRP trace $\{\tau'_1, \tau'_2, \ldots\}$.

By basic probability, we have

$$\lim_{k\to\infty} \frac{\sum_{i=1}^{k} z(\tau_i)}{k} = E[Z] \quad \text{and} \quad \lim_{k\to\infty} \frac{\sum_{i=1}^{k} z(\tau_i')}{k} = E[Z'] \, .$$

We now observe that $\tau_1, \tau_3, \tau_5, \ldots$ do not share endpoints, thus they can be regarded as truly independent and behave like a set of movements in the IRP process. The same holds for $\tau_2, \tau_4, \tau_6, \ldots$. Hence we can write

$$\begin{aligned} E[Z] &= \lim_{k\to\infty} \frac{\sum_{i=1}^{k/2} z(\tau_{2i-1})}{k} + \frac{\sum_{i=1}^{k/2} z(\tau_{2i})}{k} = \\ &= \frac{E[Z']}{2} + \frac{E[Z']}{2} = E[Z'] \, . \end{aligned}$$

If the function $z(\tau)$ is the trajectory length, the equality above implies that the expected value of the trajectory length in the RWP and IRP process is the same, i.e. $E[L] = E[L']$, where L and L' are random variables denoting the expected trajectory length in the RWP and IRP process, respectively. In the nomenclature of stochastic processes, we have thus shown a "mean–ergodic property" of the RWP mobility model, i.e., statistically there is no difference between sampling repeatedly from a single random variable L (or L') or successively from the sequence $\{L_i\}_{i\in\mathbb{N}}$. With respect to our problem, this ergodic property implies, for instance, the following: in order to determine the expected value of the trajectory length of a RWP mobile node, the analysis can be simplified by considering only the distances between two points placed uniformly at random in Q. This allows us to use the following well-known results from the theory of geometric probability (see [29]): the expected distance between two random points is $E[L] = a/3$ when the points are uniformly distributed on the one-dimensional line segment $[0, a]$, and it is $E[L] = 0.521405\, a$ when the points are uniformly distributed on the two-dimensional square $[0, a]^2$.

3.4 Components of the Node Distribution

With this formal description of RWP movement, it can be easily seen that the resulting node distribution $f_{\mathbf{X}}(\mathbf{x})$ is composed of three distinct components: the static, pause, and mobility component:

$$f_{\mathbf{X}}(\mathbf{x}) = f_s(\mathbf{x}) + f_p(\mathbf{x}) + f_m(\mathbf{x}) \, . \tag{1}$$

Although the three components of the distribution in (1) are denoted like pdfs, indeed they represent likelihood functions, i.e. their integral over Q does not necessarily corresponds to one. The static component $f_s(\mathbf{x})$ accounts for the fact that a node can remain static for the entire network operational time. The pause component $f_p(\mathbf{x})$ accounts for the time that a mobile node "rests" before starting a new movement period. Finally, the mobility component $f_m(\mathbf{x})$ accounts for the time that a mobile node is actually moving. The following two sections compute these components and finally give an equation for the overall $f_{\mathbf{X}}(\mathbf{x})$.

4 The Mobility Component of the Node Distribution

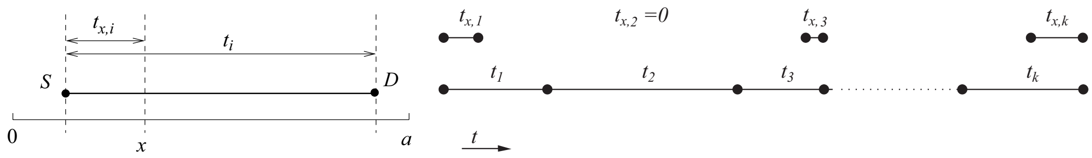
In this section, we derive the asymptotically stationary node distribution generated by the generalized RWP model under two assumptions: (a) all nodes are mobile ($p_s = 0$, no static nodes) and (b) the pause time is set to zero, i.e., $f_{T_p}(t_p) = 1$ if $t_p = 0$, and 0 otherwise. In other words, we compute the (normalized) mobility component of the overall distribution. We first consider a one-dimensional RWP model on a line segment and then extend our analysis to the two-dimensional case on a square.

4.1 One-Dimensional Case

A node moves according to the RWP model on a line segment $[0, a]$. The random variable X denotes the location of the node, where $X \in [0, a]$. Moreover, the random variables S and D denote the starting and destination points of a movement period. These points are randomly chosen from the uniform distribution on the line segment, i.e., their pdfs are

$$f_S(s) = f_D(d) = \begin{cases} \frac{1}{a} & \text{for } 0 \leq s, d \leq a \\ 0 & \text{otherwise} \end{cases} .$$

In order to derive $f_X(x)$, let us first calculate the cumulative distribution function (cdf) $F_X(x) = P(X \leq x)$, which denotes the probability that the mobile node is located within $[0, x]$ at an arbitrary instant of time. For each period i, t_i denotes the duration of this period, and $t_{x,i}$ denotes the duration that the node spends within $[0, x]$ during



(a) One movement period from S to D

(b) Movement periods and corresponding time values

Figure 1: Illustration of RWP movement on line segment $[0, a]$.

this period (see Fig. 1). If the i-th movement trajectory does not intersect $[0, x]$, we have $t_{x,i} = 0$. The corresponding random variables are denoted by T_i and $T_{x,i}$. We now observe the RWP process for a given number, say k, of movement periods. The time that the node spends in $[0, x]$ during its entire movement process ($\sum_{i=1}^{k} t_{x,i}$) divided by the total movement time of the node ($\sum_{i=1}^{k} t_i$) converges toward $P(X \leq x)$ as the number of movement periods goes to infinity:

$$P(X \leq x) = \lim_{k\to\infty} \frac{\sum_{i=1}^{k} t_{x,i}}{\sum_{i=1}^{k} t_i} = \frac{E[T_x]}{E[T]} \ .$$

In each period i, the node chooses uniformly at random a speed $v_i \in [v_{min}, v_{max}]$. Let $l_i = v_i t_i$ denote the traveled distance in period i. Similarly, let $l_{x,i} = v_i t_{x,i}$ denote the traveled distance within $[0, x]$ during this period. The corresponding random variables are denoted by V, L, and L_x, respectively. Since V and L are independent random variables, and the same holds for V and L_x, we can write $E[T] = E\left[\frac{L}{V}\right] = c \cdot E[L]$ and $E[T_x] = E\left[\frac{L_x}{V}\right] = c \cdot E[L_x]$, for some constant c that depends on the distribution of V. If V is uniformly distributed in the interval $[v_{min}, v_{max}]$, with $v_{min} > 0$, we have $c = \frac{\ln(v_{max}/v_{min})}{v_{max} - v_{min}}$ [5]. Thus, it follows immediately that

$$P(X \leq x) = \frac{E[L_x]}{E[L]} \ .$$

An important consequence of this equation is that the asymptotic cdf $F_X(x) = P(X \leq x)$ is independent of the speed choice of the nodes. As mentioned above, we have $E[L] = a/3$ from the literature on stochastic geometry. Thus, we have reduced the problem of calculating $F_X(x)$ to the problem of calculating $E[L_x]$. In order to do so, let $l_x(s, d)$ denote the value of the random variable L_x if $S = s$ and $D = d$. We have:

$$E[L_x] = \int_{s=0}^{a} \int_{d=0}^{a} l_x(s,d)\, f_S(s)\, f_D(d) \;\; dd\, ds \;\; .$$

Because of the symmetry of S and D, it is sufficient to restrict the calculation to periods with $s \leq d$ and then multiply the result by a factor of 2. A necessary condition for $l_x(s,d) \neq 0$ is that $((s \leq x) \wedge (d \leq x)) \vee ((s \leq x) \wedge (d > x))$ is true. In the first case we have $l_x(s,d) = d - s$, and in the second case we obtain $l_x(s,d) = x - s$. This yields

$$\begin{aligned} E[L_x] \;&=\; \frac{2}{a^2} \int_{s=0}^{x} \int_{d=s}^{x} (d-s)\, dd\, ds \, + \\ &+ \frac{2}{a^2} \int_{s=0}^{x} \int_{d=x}^{a} (x-s)\, dd\, ds \;=\; -\frac{2}{3a^2}\, x^3 + \frac{1}{a}\, x^2 \;\; . \end{aligned}$$

The cdf of X is therefore given by

$$F_X(x) = \frac{E[L_x]}{E[L]} = -\frac{2}{a^3}\, x^3 + \frac{3}{a^2}\, x^2 \; , \quad \text{for } \; 0 \leq x \leq a \, .$$

The probability of finding a node between x_1 and x_2 is $P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$. For example, a node is expected to reside 68.75% of its movement time within $\left]\frac{a}{4}, \frac{3a}{4}\right]$, i.e., within the central 50% of the line segment. Using the definition of pdf $f_X(x) = \frac{\partial F_X(x)}{\partial x}$, we can conclude with the following result.

Theorem 1 *The asymptotically stationary pdf of the location X of a mobile node moving on a line segment $[0,a]$ according to the generalized RWP model with $p_s = 0$ and $t_p = 0$ is*

$$f_X(x) = -\frac{6}{a^3}\, x^2 + \frac{6}{a^2}\, x$$

for $0 < x < a$, and 0 otherwise. Furthermore, the asymptotic distribution is independent of the value of v_{max} and $v_{min} > 0$ and the initial node distribution.

This function represents the normalized version of the mobility component $f_m(x)$ of the overall node distribution (1). It is illustrated in Figure 2 and has been validated by simulations. The probability of finding a node close to the border of the line segment goes to zero; the maximum value of $f_X(x)$ is at $x = 0.5a$, and the expected location of a node is $E[X] = 0.5\, a$.

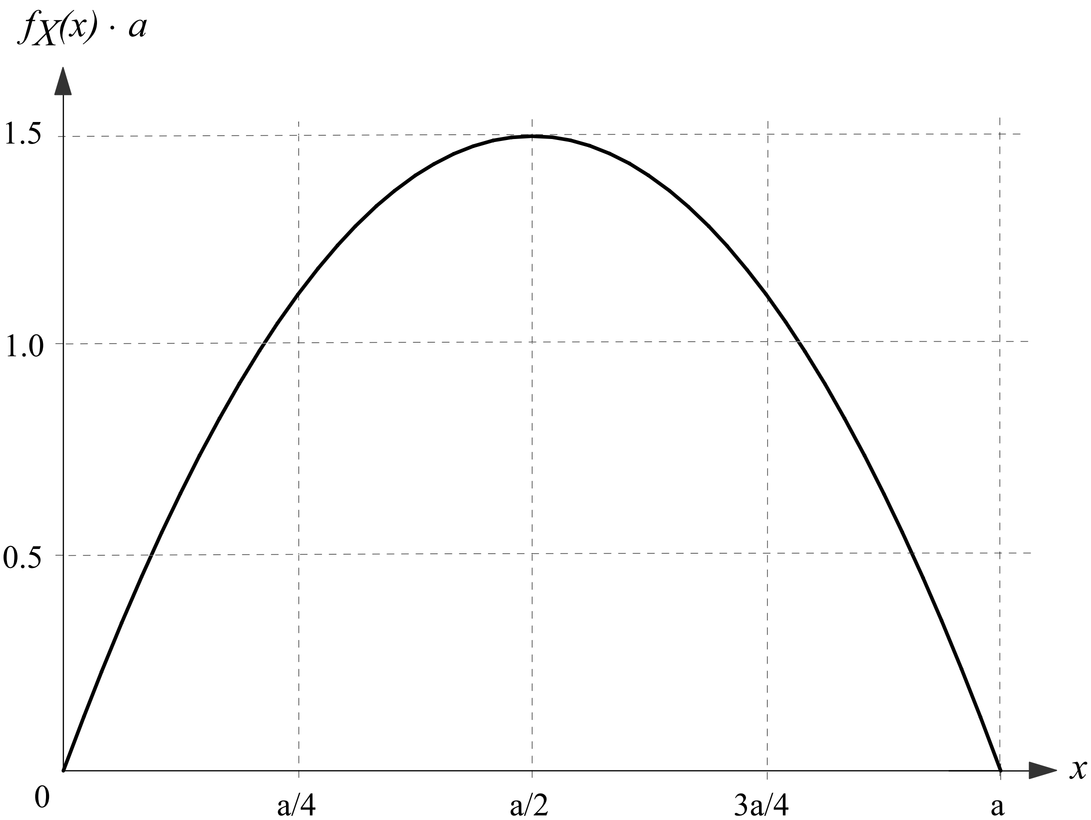


Figure 2: The asymptotic pdf $f_X(x)$ of RWP movement on a line segment

4.2 Two-Dimensional Case

In this section we consider the mobility component ($p_s = t_p = 0$) of the spatial distribution in a two-dimensional unit square $Q = [0, a]^2$. For simplicity, we set $a = 1$. The Cartesian coordinates of a mobile node are $\mathbf{X} = (X, Y)$, where $X, Y \in [0, 1]$. The asymptotic distribution is denoted by $f_{\mathbf{X}}(\mathbf{x}) = f_{XY}(x, y)$. The starting and destination points, denoted by $\mathbf{D} = (D_x, D_y)$ and $\mathbf{S} = (S_x, S_y)$, are uniformly distributed in Q. Specific values of the random variables are denoted by $\mathbf{x}, x, y, \mathbf{d}, d_x, d_y$, and so on.

First of all, we note that the distribution in two dimensions cannot be directly derived from the equation of the one-dimensional case. In fact, the two-dimensional movement is composed of two *dependent* one-dimensional movements. The speed of a node projected along the x-axis is not constant in general, and it is different from the (non-constant) speed along the y-axis. As we have shown in [2] and [26], the simple product

$$f_X(x)\, f_X(y) = 36\, xy\, (x-1)\, (y-1) \;\;, \text{ for } 0 \le x, y \le 1 \;,$$

yields an approximation of the distribution $f_{XY}(x,y)$. Nevertheless, there is a non-negligible difference between $f_X(x)\,f_X(y)$ and $f_{XY}(x,y)$, and this is why we are interested in a better expression for the distribution.

To derive the exact expression of $f_{XY}(x,y)$, we could use the same technique as in the one dimensional case, i.e., calculate $F_{XY}(x,y) = P((X<x)\wedge(Y<y))$ and then differentiate. However, the integration of the length $l_{xy}(\mathbf{s},\mathbf{d})$, i.e., the moved distance within the region defined by $\{(X,Y)\in Q \mid (X\leq x)\wedge(Y\leq y)\}$, over all possible starting and ending points is very difficult.

For this reason, we use a different technique, which directly computes a very good approximation of $f_{XY}(x,y)$. Let us assume for a moment that the node velocity is constant during the entire observation period, i.e., $v_{min}=v_{max}=v>0$. With this assumption, we can refer to the length and duration of a trajectory interchangeably. Let

$$\begin{aligned} P(x,y,\delta) \;&=\; P\left((x-\tfrac{\delta}{2}<X\leq x+\tfrac{\delta}{2})\wedge(y-\tfrac{\delta}{2}<Y\leq y+\tfrac{\delta}{2})\right) \\ &=\; \int_{x-\delta/2}^{x+\delta/2}\int_{y-\delta/2}^{y+\delta/2} f_{XY}(x_0,y_0)\;dy_0\;dx_0 \end{aligned}$$

denote the probability that the node is in a square of length δ centered in $\mathbf{x}=(x,y)$. This square is denoted as Q_{xy}^{δ} in the following (see Fig. 3). If δ is sufficiently small, $f_{XY}(x,y)$ can be considered to be constant in Q_{xy}^{δ}, and $P(x,y,\delta)$ can be rewritten as $P(x,y,\delta)=\delta^2 f_{XY}(x,y)$. This yields

$$f_{XY}(x,y)=\lim_{\delta\to 0}\frac{P(x,y,\delta)}{\delta^2}\;.$$

We now consider a fixed square Q_{xy}^{δ} positioned at $\mathbf{x}=(x,y)$, and a trajectory $\tau(\mathbf{s},\mathbf{d})$ between $\mathbf{s}$ and $\mathbf{d}$. As illustrated in Fig. 3, l denotes the total length of the trajectory, i.e., $l=l(\mathbf{s},\mathbf{d})=\|\tau(\mathbf{s},\mathbf{d})\|$, and l_{xy}^{δ} the sub-length inside Q_{xy}^{δ}, i.e., $l_{xy}^{\delta}=l_{xy}^{\delta}(\mathbf{x},\mathbf{s},\mathbf{d},\delta)=\|\tau(\mathbf{s},\mathbf{d})\cap Q_{xy}^{\delta}\|$. The corresponding random variables are denoted by L and L_{xy}^{δ}, respectively. Clearly, $l_{xy}^{\delta}=0$ for all $\tau(\mathbf{s},\mathbf{d})$ that do not intersect Q_{xy}^{δ}. As in the one-dimensional case, we can define the expected sub-length $E[L_{xy}^{\delta}]$ of a random trajectory inside a given Q_{xy}^{δ}, and write

$$P(x,y,\delta)=\frac{E[L_{xy}^{\delta}]}{E[L]}\;. \tag{2}$$

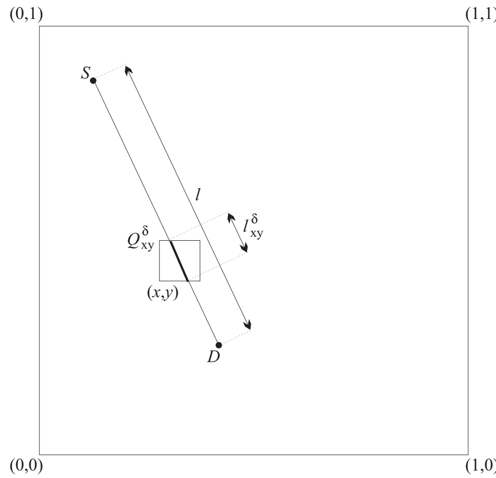


Figure 3: Intersection of a trajectory $\tau(\mathbf{s}, \mathbf{d})$ with Q^{δ}_{xy}.

The expected trajectory length $E[L]$ of the RWP model is equivalent to the expected distance between two independent points chosen uniformly at random in $Q = [0,1]^2$, which is $E[L] = 0.521405$ [29].

The expected value $E[L^{\delta}_{xy}]$ depends on the side δ and on the position $\mathbf{x}$ of the small square Q^{δ}_{xy}, and can be calculated as the integral of $l^{\delta}_{xy}(\mathbf{x}, \mathbf{s}, \mathbf{d}, \delta)$ over all possible starting and destination points in Q, i.e.,

$$E[L^{\delta}_{xy}] = \int_{\mathbf{s}\in Q} \int_{\mathbf{d}\in Q} l^{\delta}_{xy}(\mathbf{x}, \mathbf{s}, \mathbf{d}, \delta) f_{\mathbf{S}}(\mathbf{s})\, f_{\mathbf{D}}(\mathbf{d})\, d\mathbf{d}\, d\mathbf{s}\,, \tag{3}$$

where $f_{\mathbf{D}}(\mathbf{d}) = f_{\mathbf{S}}(\mathbf{s}) = 1$ for $\mathbf{s}, \mathbf{d} \in [0,1]^2$, and $d\mathbf{d} = dd_x dd_y$ as well as $d\mathbf{s} = ds_x ds_y$.

Let us first consider the inner integral for a fixed starting point $\mathbf{s}$. Only destination points $\mathbf{d}$ for which the trajectory $\tau(\mathbf{s}, \mathbf{d})$ intersects Q^{δ}_{xy} contribute to the integral. This is illustrated in Figure 4: for given $\mathbf{s}$, only destination points inside the shaded polygon yield $l^{\delta}_{xy} \neq 0$. Denoting this polygon by $A(\mathbf{x}, \mathbf{s}, \delta)$, we can state

$$\int_{\mathbf{d}\in Q} l^{\delta}_{xy}(\mathbf{x}, \mathbf{s}, \mathbf{d}, \delta)\, d\mathbf{d} = \int_{\mathbf{d}\in A(\mathbf{x},\mathbf{s},\delta)} l^{\delta}_{xy}(\mathbf{x}, \mathbf{s}, \mathbf{d}, \delta)\, d\mathbf{d}\,.$$

Determining the exact expression of this integral seems to be very difficult. In the following, we conjecture that

$$\begin{aligned} \int_{\mathbf{d}\in A(\mathbf{x},\mathbf{s},\delta)} l^{\delta}_{xy}(\mathbf{x}, \mathbf{s}, \mathbf{d}, \delta)\, d\mathbf{d} &\approx \int_{\mathbf{d}\in A(\mathbf{x},\mathbf{s},\delta)} c_1 \delta\, d\mathbf{d} \\ &= c_1\, \delta\, A(\mathbf{x}, \mathbf{s}, \delta)\,, \end{aligned}$$

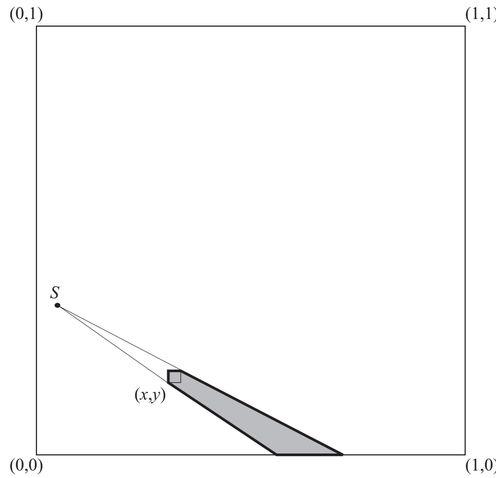


Figure 4: The area of the shaded polygon, denoted as $A(\mathbf{x}, \mathbf{s}, \delta)$, represents the probability that a trajectory that starts at $\mathbf{s}$ intersects Q^{δ}_{xy}.

for some constant $c_1 > 0$; i.e., independently of the coordinates of $\mathbf{x}$, $\mathbf{s}$, and $\mathbf{d}$, the function $l^{\delta}_{xy}(\mathbf{x}, \mathbf{s}, \mathbf{d}, \delta)$ is accurately approximated by $c_1\delta$. The validity of this conjecture is confirmed by the experimental analysis reported in Section 6. If the conjecture holds, we can rewrite (3) as

$$E[L^{\delta}_{xy}] \approx c_1 \delta \int_{\mathbf{s} \in Q} A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s}\ .$$

The area of the polygon $A(\mathbf{x}, \mathbf{s}, \delta)$, divided by the total area (which is 1), represents the probability that a trajectory intersects Q^{δ}_{xy} under the condition that this trajectory starts at $\mathbf{s}$. The probability that a random trajectory intersects Q^{δ}_{xy} can thus be calculated as the integral of $A(\mathbf{x}, \mathbf{s}, \delta)$ over all possible positions of $\mathbf{s}$ in the deployment region Q. Let this probability be denoted by

$$P_{\tau}(x, y, \delta) = \int_{\mathbf{s} \in Q} A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s}. \tag{4}$$

Plugging the above two equations into (2), we can write:

$$f_{XY}(x, y) = \lim_{\delta \to 0} \frac{P(x, y, \delta)}{\delta^2} \approx c \cdot \lim_{\delta \to 0} \frac{P_{\tau}(x, y, \delta)}{\delta}\ .$$

Up to a constant $c = c_1/E[L] > 0$ and an approximation, we have reduced the original problem to the problem of determining the probability that a random trajectory intersects Q^{δ}_{xy}. Observe that it is not necessary to calculate the value of the constant c, since it will be absorbed by the multiplicative constant needed to normalize $f_{XY}(x, y)$.

Finding the exact expression of the area $A(\mathbf{x}, \mathbf{s}, \delta)$ is not straightforward. The shape of the polygon depends on the positions $\mathbf{s}$ and $\mathbf{x}$. For this reason, given the coordinate $\mathbf{x}$, we divide Q into a number of subareas, with the property that all the starting points in the same subarea induce polygons with the same shape. This way, we can calculate the partial integral independently on each subarea, and obtain the overall integral as the sum of the contributes of all the subareas. Details on how $A(\mathbf{x}, \mathbf{s}, \delta)$ and, consequently, $f_{XY}(x, y)$ are calculated can be found in the Appendix. In summary, we obtain the following result.

Theorem 2 *The asymptotically stationary pdf of the location $\mathbf{X} = (X, Y)$ of mobile nodes moving in $[0,1]^2$ according to the generalized RWP model with $p_s = 0$ and $t_p = 0$ can be closely approximated by*

$$
\begin{aligned}
f_{XY}(x,y) &= \\
&= \begin{cases}
f^*_{XY}(x,y) & 0 < x \le \frac{1}{2},\ 0 < y \le x \\
f^*_{XY}(y,x) & 0 < x \le \frac{1}{2},\ x \le y \le \frac{1}{2} \\
f^*_{XY}(1-y,x) & 0 < x \le \frac{1}{2},\ \frac{1}{2} \le y \le 1-x \\
f^*_{XY}(x,1-y) & 0 < x \le \frac{1}{2},\ 1-x < y \le 1 \\
f^*_{XY}(1-x,y) & \frac{1}{2} \le x < 1,\ 0 < y \le 1-x \\
f^*_{XY}(y,1-x) & \frac{1}{2} \le x < 1,\ 1-x \le y \le \frac{1}{2} \\
f^*_{XY}(1-y,1-x) & \frac{1}{2} \le x < 1,\ \frac{1}{2} \le y \le x \\
f^*_{XY}(1-x,1-y) & \frac{1}{2} \le x < 1,\ x \le y < 1 \\
0 & \text{otherwise}
\end{cases},
\end{aligned}
$$

where f^*_{XY} is defined on $Q^* = \{(x,y) \in [0,1]^2 \mid (0 < x \le 0.5) \wedge (0 < y \le x)\}$, with

$$
\begin{aligned}
f^*_{XY}(x,y) &= 6y + \frac{3}{4}\left(1 - 2x + 2x^2\right)\left(\frac{y}{y-1} + \frac{y^2}{(x-1)\,x}\right) \\
&+ \frac{3y}{2}\Bigg[(2x-1)(y+1)\ln\left(\frac{1-x}{x}\right) \\
&\quad + \left(1 - 2x + 2x^2 + y\right)\ln\left(\frac{1-y}{y}\right)\Bigg].
\end{aligned}
$$

Again, $f_{XY}(x, y)$ corresponds to the normalized mobility component $f_m(\mathbf{x})$. Its plot and some contour lines are shown in Figure 5. As in the one-dimensional case, the expected

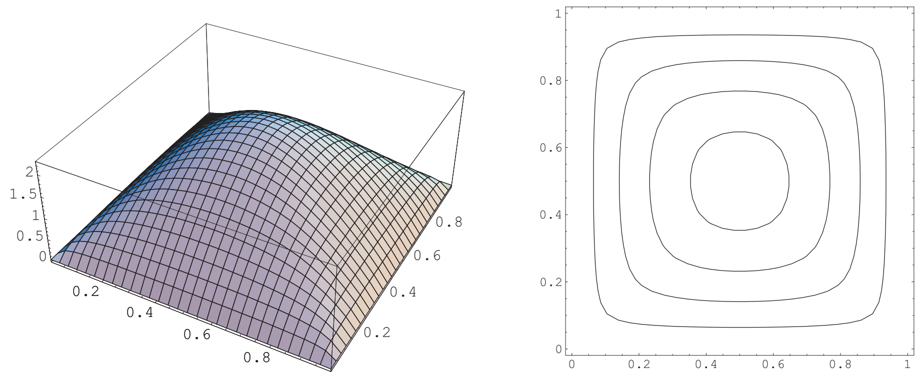


Figure 5: Plot of the mobility component and contour lines corresponding to the values $f_m(x, y) = 0.5$, 1, 1.5 and 2.

location of a node and the maximum of the density are in the middle of the region, at $\mathbf{x} = (0.5, 0.5)$. While the density in the middle is rotary symmetric, the contour lines toward the border become more and more rectangular. The probability of finding a node at the borders of the region goes to zero. Note that, as in the one-dimensional case, the asymptotic distribution of mobile nodes is independent of the initial node distribution. Furthermore, the proof that $f_{\mathbf{X}}(\mathbf{x})$ in one-dimensional networks is independent of the choice of the node velocities can be generalized to the two-dimensional setting.

5 Node Distribution of the Generalized RWP model

In this section, we first analyze the static component $f_s(\mathbf{x})$ and the pause component $f_p(\mathbf{x})$, then perform proper scaling of the mobility component $f_m(\mathbf{x})$, and finally show the equation of the overall distribution $f_{\mathbf{X}}(\mathbf{x}) = f_s(\mathbf{x}) + f_p(\mathbf{x}) + f_m(\mathbf{x})$.

The static component $f_s(\mathbf{x})$ can be determined in a straightforward manner from the initial distribution by observing that a node remains static with probability p_s. Thus, we have

$$f_s(\mathbf{x}) = p_s\, f_{init}(\mathbf{x}) \ ,$$

independently of the time t at which the node is observed.

Let us now consider a node that is not static. During its RWP movement, it alternates between pause periods (lasting $t_{p,i}$ for the i-th period) and movement periods (lasting

t_i). The pause periods contribute to the pause component, and the movement periods contribute to the mobility component.

For derivation of the pause component, we define $p_p(t)$ as the probability that a node is pausing at time t. Since the destination points are uniformly distributed in $[0,1]^\alpha$, we write

$$f_p(\mathbf{x},t) = (1-p_s)\, p_p(t)$$

for $\mathbf{x} \in [0,1]^\alpha$, and 0 otherwise. Since we are interested in characterizing the asymptotic density, we must determine the value of $p_p(t)$ for $t \to \infty$. Assuming that $v_{min} = v_{max} = v > 0$, the duration of a movement period depends only on the distance between the starting and destination waypoint, i.e., $t_i = \frac{l_i}{v}$. The total running time of the RWP process after period k is given by

$$\sum_{i=1}^{k} \left(t_{p,i} + \frac{1}{v}\, l_i \right) .$$

The probability that a node is resting at a randomly chosen time instant therefore is

$$p_p = \lim_{k\to\infty} \frac{\sum_{i=1}^{k} t_{p,i}}{\sum_{i=1}^{k} t_{p,i} + \frac{1}{v}\sum_{i=1}^{k} l_i} = \frac{E[T_p]}{E[T_p] + \frac{1}{v}E[L]} ,$$

where $E[T_p]$ is the expected value of the pause time distribution $f_{T_p}(t_p)$, and $E[L]$ is the well-known expected trajectory length.

In order to finally obtain the mobility component, the results on the node distribution of the previous sections must be scaled, taking into account the probability that a node is actually moving at an arbitrary time. Denoting by $f_m^0(\mathbf{x})$ the distribution of Theorem 1 and 2, the mobility component is given by

$$f_m(\mathbf{x}) = (1-p_s)(1-p_p)\, f_m^0(\mathbf{x}) .$$

Knowing all three components, we are now ready to present the main result of this paper.

Theorem 3 *The asymptotically stationary pdf of the location* $\mathbf{X}$ *of nodes moving in* $Q = [0,1]^\alpha$*, with* $\alpha = 1,2$*, according to the generalized RWP model with constant velocity* $v > 0$*, is*

$$f_{\mathbf{X}}(\mathbf{x}) = p_s\, f_{init}(\mathbf{x}) + (1-p_s)\, p_p + (1-p_s)(1-p_p) f_m^0(\mathbf{x})$$

for $\mathbf{x} \in [0,1]^\alpha$*, and* 0 *otherwise, where* $p_p = \left(1 + \frac{E[L]}{v\,E[T_p]}\right)^{-1}$*, with* $E[L] = 1/3$ *for* $\alpha = 1$ *and* $E[L] = 0.521405$ *for* $\alpha = 2$*. The normalized mobility component* $f_m^0(\mathbf{x})$ *is defined in Theorem 1* ($\alpha = 1$) *and Theorem 2* ($\alpha = 2$)*, respectively.*

Let us discuss this result, under the assumption that the pause time is fixed to an arbitrary value t_p. If the initial node distribution $f_{init}(\mathbf{x})$ is uniform, the asymptotic node distribution $f_{\mathbf{X}}(\mathbf{x})$ is the sum of a uniform and a non-uniform component. As p_s and/or t_p increase, the uniform component of the density becomes predominant, and $f_{\mathbf{X}}(\mathbf{x})$ can be approximated with the uniform distribution. Conversely, for small values of p_s and/or t_p, the non-uniform component dominates and generates a significant "border effect." These observations are fully coherent with the statistical analysis presented in [8]. The influence of the velocity v on $f_{\mathbf{X}}(\mathbf{x})$ is less evident: in general, higher velocities cause a shorter movement duration and, consequently, a "more uniform" distribution. For extreme values of t_p the effect of the velocity is negligible. If $t_p = 0$, the pause component of the density will be zero (regardless of the value of v) and the density $f_{\mathbf{X}}(\mathbf{x})$ will be independent of v. Similarly, if t_p is very large, then $p_p \approx 1$ regardless of the value of v.

We remark that Theorem 3 has practical relevance for simulation-based studies of RWP mobile networks. So far, the only way to investigate relevant asymptotic properties of mobile networks was to simulate the nodes movement for a very large number of steps. This is done at the expense of considerable computational resources. As a consequence, the number of nodes in the mobile system is usually kept small (it is rarely above 100 in existing experimental results). As wireless ad hoc networks will become reality in a near future, their size is likely to grow to as much as thousands of nodes. Hence, the simulation of *large* mobile networks, in which the scalability of the protocols can be carefully investigated, will become an issue. We believe that our characterization of the node spatial distribution of mobile networks is of great help in the simulation of large mobile ad hoc networks.

The generation of nodes' positions can be done as follows. A node remains static during the entire simulation with probability p_s. If the node is non–mobile, its position is chosen according to the initial distribution $f_{init}(\mathbf{x})$. If the node is mobile, its position is chosen according to the equation of Theorem 3, where p_s is set to zero[1] and the other parameters reflect the settings of the RWP mobility parameters in the simulated scenario. It can be easily seen that this procedure puts the system immediately into its asymptotic "steady state," thus avoiding the large number of movement periods needed to make the system converge to this state.

6 Experimental Evaluation

In this section we report the results of the simulation-based experiments that we have performed to evaluate how well the equations for the two-dimensional node distribution approximate the actual distribution.

Our simulation tool takes as input the mobility parameters of the RWP model: the probability p_s, the pause time t_p, and the parameters of the node velocity (v_{min}, v_{max}). In the remainder, time measures are expressed as the number of time steps, and length and velocity measures are normalized with respect to the unit square. A number of $n = 1000$ nodes are distributed uniformly and independently at random in $[0,1]^2$; then, they start moving according to the RWP mobility model. Later on, we will show simulation results where the initial node distribution is not uniform, and the pause time and velocity are randomly chosen.

In order to record the node spatial distribution, we divide $[0,1]^2$ into a number of square cells of the same size, arranged in a grid fashion. In our experiments, we use a grid of 31×31 cells with side lengths $1/31$. After $t = 10\,000$ time steps, the number of nodes in each cell is recorded. These numbers are accumulated over 10 000 simulation runs, and are reported as the result of the experiment. These values for the number of mobility steps and simulation runs are chosen because they are a good compromise between statistical accuracy and running time. If the theoretical analysis is accurate, the normalized plot obtained by using these data should closely resemble that obtained by our equation.

In the first series of experiments, we consider a scenario with mobile nodes only and

[1] Since static nodes are considered separately in the position generation process, the static component of the density $f(\mathbf{x})$ must be set to 0.

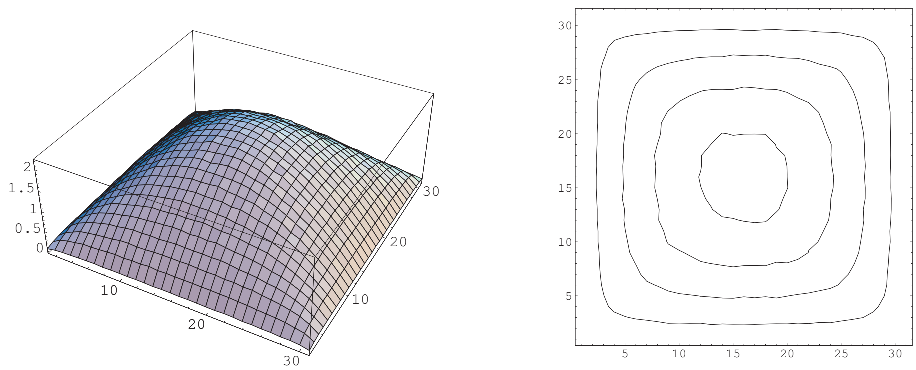
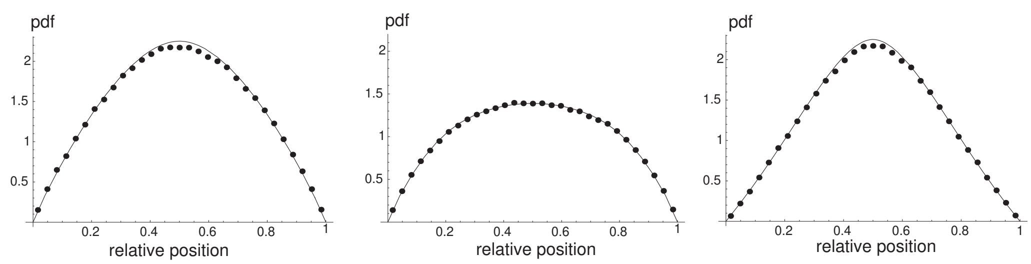
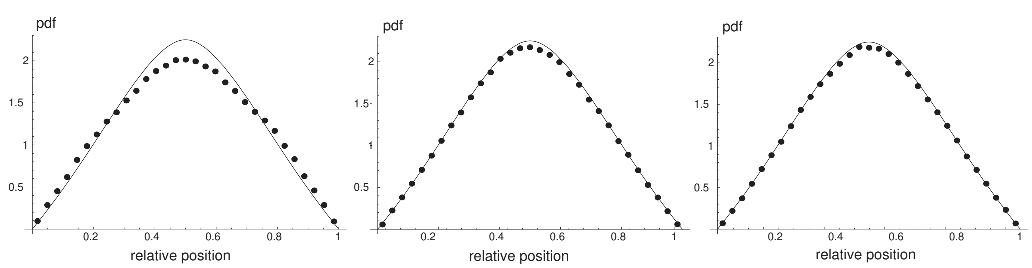


Figure 6: Node distribution obtained by simulation for $p_s = t_p = 0$ (mobility component). The z-axis reports the number of times a node is "observed" in the given cell (after normalization). The contour lines on the right correspond to the values $f_{XY}(x, y) = 0.5$, 1, 1.5, and 2.



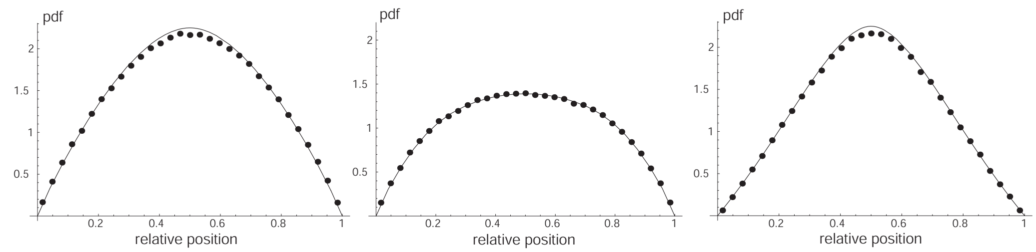
a) Cut at $y = 0.5$ b) Cut at $y \approx 0.21$ c) Diagonal cut

Figure 7: Experiment 1: Node distribution for $p_s = t_p = 0$.



a) $t = 100$ steps b) $t = 500$ steps c) $t = 1000$ steps

Figure 8: Experiment 2: Diagonal cut of the node distribution for $p_s = t_p = 0$ (mobility component) and $v = 0.01$ after 100, 500, and 1000 steps.



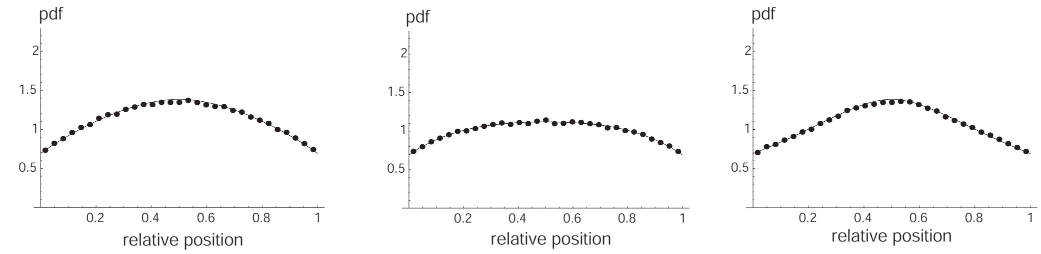
a) Cut at $y = 0.5$ b) Cut at $y \approx 0.21$ c) Diagonal cut

Figure 9: Experiment 3: Node distribution for $p_s = t_p = 0$ (mobility component), with velocity chosen uniformly at random in the interval $[0.005, 0.015]$.

zero pause time. Our goal is to evaluate the impact of the approximation that we made in the derivation of the two-dimensional mobility component (Theorem 2). We set $p_s = 0$, $t_p = 0$, and $v_{min} = v_{max} = v = 0.01$. The normalized plot of the recorded node distribution over (x, y) and the corresponding contour lines are reported in Figure 6. They show a very close resemblance with the plots of the theoretically derived function of $f_m(x, y)$ in Figure 5. This resemblance is further evidenced by the plots shown in Figure 7. These graphics report two cuts parallel to the x-axis (for $y = 0.5$ and $y \approx 0.21$) and the diagonal cut of the 3D plot. Experimental data are represented by bold points, and the lines show the theoretical curves. The result of this experiment shows that the approximation that we made in the derivation of $f_m(x, y)$ does not significantly affect the quality of the result.

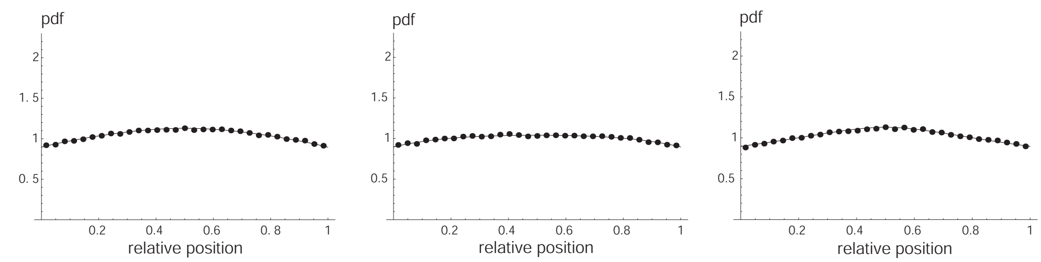
In a second experiment, we evaluate the rate of convergence of the node distribution to the asymptotic distribution. Since the mobility component of the distribution is the most critical from this point of view, we set $p_s = t_p = 0$. The diagonal cut of the node distribution resulting after $t = 100$, 500, and 1000 steps with $v = 0.01$ is shown in Figure 8. As seen from the figure, $t = 500$ steps seem to be sufficient to achieve the asymptotic distribution. The expected number of movement periods during t time steps is given by $\frac{t}{E[T]} = \frac{t\,v}{E[L]} = \frac{t \cdot 0.01}{0.521405}$. Thus, as a rule of thumb, we can say that about $\frac{500 \cdot 0.01}{0.521405} \approx 10$ movement periods are on average needed to achieve the stationary distribution.

In the third experiment, we validate that the normalized mobility component of the distribution is actually independent of the choice of the velocity. To this purpose, we set $p_s = t_p = 0$ as in the previous experiment, while the node velocities are chosen uniformly



a) Cut at $y = 0.5$ b) Cut at $y \approx 0.21$ c) Diagonal cut

Figure 10: Experiment 4.1: Node distribution for $p_s = 0.1$, $t_p = 100$, and $v = 0.01$.



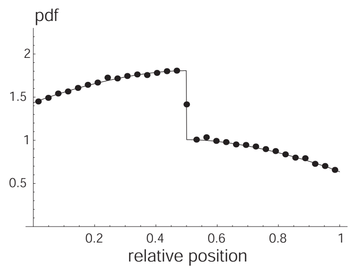
a) Cut at $y = 0.5$ b) Cut at $y \approx 0.21$ c) Diagonal cut

Figure 11: Experiment 4.2: Node distribution for $p_s = 0.3$, $t_p = 300$, and $v = 0.01$.

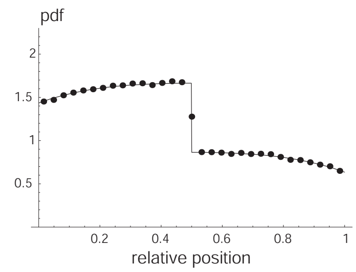
at random in the interval $[0.005, 0.015]$. The results of this experiment are reported in Figure 9. As it can be seen, the effect of allowing randomly chosen values for the node velocities is negligible. This fact confirms that Theorem 2 holds also when the velocity is chosen uniformly at random in the interval $[v_{min}, v_{max}]$.

In the fourth series of experiments, we study how well our equation of the complete node distribution (Theorem 3) fits the experimental data in two hybrid scenarios. In a first scenario, we set $p_s = 0.1$, $t_p = 100$, and $v = 0.01$; and in a second scenario, we set $p_s = 0.3$, $t_p = 300$, and $v = 0.01$. The complete node distribution is now composed of two uniform distributions (the static and pause component) and a non-uniform scaled mobility component. In the first scenario, we have $f_{\mathbf{X}}(\mathbf{x}) = 0.1 + 0.9 \cdot 0.6573 + 0.3084 \cdot f_m(\mathbf{x}) = 0.692 + 0.308 \cdot f_m(\mathbf{x})$. The second scenario yields $f(\mathbf{x}) = 0.896 + 0.104 \cdot f_m(\mathbf{x})$. The results of the experiments are reported in Figures 10 and 11. As it can be seen, our equation is a very good fit of the experimental data in both scenarios.

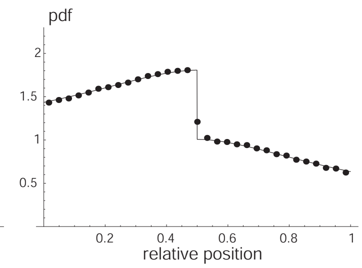
Finally, we verify the quality of our equation for the generalized RWP model with non-uniform initial distribution and random pause times. For this purpose, we extended the



a) Cut at $y = 0.5$



b) Cut at $y \approx 0.21$



c) Diagonal cut

Figure 12: Experiment 5: Node distribution for $p_s = 0.2$, $f_{init} = 4$ for $Q_0 = [0, 0.5]^2$ and 0 otherwise, $f_{T_p}(t_p) = 1/200$ for $100 \leq t_p \leq 300$ and 0 otherwise, and velocity chosen uniformly at random in $[0.005, 0.015]$.

simulator by allowing nodes to be initially distributed uniformly at random in the subarea $[0, 0.5]^2$. Further, the pause time and the velocity are chosen uniformly at random between a minimum and maximum value for each movement. We simulate the following scenario: nodes initially distributed uniformly at random in $[0, 0.5]^2$, $p_s = 0.2$, the pause time is chosen uniformly at random in the interval $[100, 300]$ at each movement (independently for each node), and v is taken from $[0.005, 0.015]$. The result of the experiment is shown in Figure 12. Also in this case, the equation fits the experimental data very well.

7 Conclusions

The theoretical results presented in this paper have significant practical relevance. First, they allow us to *improve the simulation methodology* used in the ad hoc networking research community. By initially distributing the nodes according to our distribution $f_{\mathbf{X}}(\mathbf{x})$, we put the network in its asymptotic "steady state," thus avoiding the number of movement periods needed to make the system converge to this state. Thus, the computational resources can be used to investigate the behavior of the network *after* the steady state has been reached, rather than wasted in investigating the startup phase. From this point of view, our work can be seen as complementary to a recent paper by Yoon et al. [34], paving the way towards more accurate simulation of ad hoc networks.

Second, our results serve as a starting point for the *analytical investigation* of ad hoc networks with RWP mobility. Given the distribution $f_{\mathbf{X}}(\mathbf{x})$ derived in this paper, the average route length and the connectivity (just to cite two well studied properties in the static case) in presence of RWP mobility can be analyzed in a theoretical framework. For example, we can now compute the simulation parameters needed to obtain an almost surely connected ad hoc network with RWP mobility. By setting these parameters accordingly, a researcher can be sure that the simulated RWP network is connected during most of the simulation time. A first step forward in this direction was made in [4]. Note that, without an explicit expression for $f_{\mathbf{X}}(\mathbf{x})$, it is impossible to compare simulation results based on RWP mobility with analytical results, because the latter are typically based on uniform node distribution.

Last but not least, the derivation of $f_{\mathbf{X}}(\mathbf{x})$ gives us a better understanding on how the RWP model behaves and why it behaves like this. For example, we have verified that the asymptotically stationary normalized mobility component is independent of the speed choice of the nodes and their initial spatial distribution.

A Calculating f_{XY}

The division of the unit square into subareas for a given position of Q^{δ}_{xy} is reported in Figure 13. First, we divide $Q = [0,1]^2$ into four quadrants $Q_1, \ldots, Q_4$. Quadrants are separated from each other by strips of width δ, obtained by extending the sides of Q^{δ}_{xy} to the borders. Each quadrant is then further divided into three sub-quadrants, obtained by extending the lines that connect the opposite corner of the unit square to opposite vertices of Q^{δ}_{xy}. We then have a total of 16 regions. For clarity, only the division of the first quadrant is shown in Fig. 13. We observe that the area of some of these regions approaches 0 as $\delta \to 0$, hence their contribution can be omitted when calculating the value of the overall integral. This is the case of the area of the four strips of width δ, as well as of the area of Q_{12} and of the corresponding regions in the other quadrants. Thus, we can rewrite the overall integral of (4) as

$$\iint_Q A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s} = \sum_{i=1,\ldots,4} \iint_{Q_i} A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s}\ ,$$

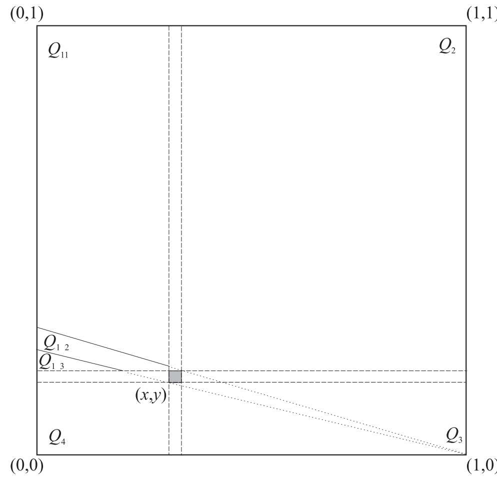


Figure 13: Division of unit square into quadrants and sub-quadrants.

where each of the summands can be computed by

$$\iint_{Q_i} A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s} \ = \iint_{Q_{i1}} A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s} + \iint_{Q_{i3}} A(\mathbf{x}, \mathbf{s}, \delta)\ d\mathbf{s}\ . \tag{5}$$

In the remaining derivation of $f_{XY}(x, y)$, we make use of symmetries in the problem. Since we are considering a square deployment region, it is sufficient to know the node distribution in the subregion $Q^* = \{(x, y) \in [0, 1]^2 \,|\, (0 < x \leq 0.5) \wedge (0 < y \leq x)\}$. The distribution in the other subregions of Q is obtained by proper variable substitutions. We denote the distribution in Q^* by $f^*_{XY}(x, y)$.

Under the assumption that Q^δ_{xy} is located in Q^*, let us detail the calculation of one summand in (5). In general, the area of a convex polygon defined by n points $\mathbf{x}_1 = (x_1, y_1), \ldots, \mathbf{x}_n = (x_n, y_n)$ can be calculated by $A = \frac{1}{2}((x_1 - x_2)(y_1 + y_2) + (x_2 - x_3)(y_2 + y_3) + \ldots + (x_n - x_1)(y_n + y_1))$, where the n points must be ordered counterclockwise. Let

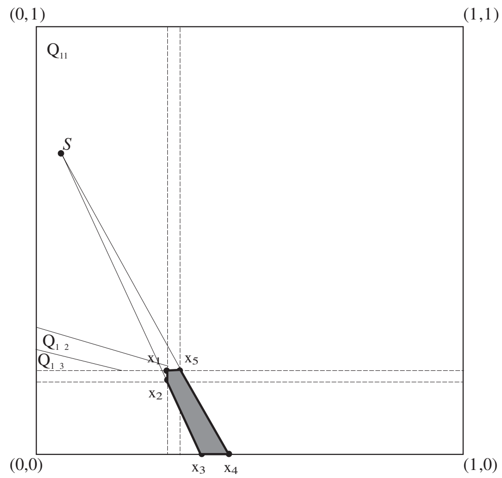


Figure 14: Area $A(\mathbf{x}, \mathbf{s}, \delta)$ if $\mathbf{s}$ is in Q_{11}.

us assume that $\mathbf{s}$ is in Q_{11}. We have (see Fig. 14)

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \\ x_5 & y_5 \end{pmatrix} = \begin{pmatrix} x - \frac{\delta}{2} & y + \frac{\delta}{2} \\ x - \frac{\delta}{2} & y - \frac{\delta}{2} \\ \frac{(x-\delta/2-s_x)s_y}{s_y-y+\delta/2} + s_x & 0 \\ \frac{(x+\delta/2-s_x)s_y}{s_y-y-\delta/2} + s_x & 0 \\ x + \frac{\delta}{2} & y + \frac{\delta}{2} \end{pmatrix} ,$$

which yields

$$A(\mathbf{x}, \mathbf{s}, \delta) = \\ = \frac{1}{2}\,\delta\left(\delta - x + s_x + y - s_y + \frac{4s_y^2(s_x-x+y-s_y)}{(2y-2s_y+\delta)(2s_y-2y+\delta)}\right) .$$

This area must be integrated over Q_{11}. Observing that the line that delimits the lower side of Q_{11} has equation $y(x) = m_{11}x + q_{11}$, where $m_{11} = \frac{2y+\delta}{2x-2+\delta}$ and $q_{11} = -m_{11}$, we write

$$\iint_{Q_{11}} A(\mathbf{x}, \mathbf{s}, \delta)\; d\mathbf{s} = \int_0^{x-\frac{\delta}{2}} \int_{m_{11}s_x+q_{11}}^{1} A(\mathbf{x}, s_x, s_y, \delta)\; ds_y\; ds_x \,. \tag{6}$$

Taking the limit of (6), divided by δ, as δ goes to 0, yields

$$\lim_{\delta \to 0} \frac{(6)}{\delta} = \frac{1}{4}\, xy \left[\frac{x^2 + x(6y-7) + 2(3-4y+y^2)}{(x-1)(y-1)} \right.$$
$$\left. + 2(x+y)\ln\left[\left(\frac{1}{x}-1\right)\left(\frac{1}{y}-1\right)\right]\right]$$

This term gives the contribution of Q_{11} to the node density. The derivation of the partial integrals referring to the other regions can be obtained by similar geometric arguments, and is not reported for the sake of brevity. The overall density can be calculated by summing the contribution of all the regions.

The resulting expression, which we denote $f^*_{XY}(x, y)$, must be normalized in such a way that $\int_{Q^*} f^*_{XY}(x, y)dxdy = \frac{1}{8}$, since the area of Q^* is $\frac{1}{8}$. After long and tedious calculation, which is not reported, we have obtained the expression reported in the statement of Theorem 2.

Acknowledgments

C. Bettstetter was supported by the German Science Foundation (DFG). Some parts of this article are based on the authors' workshop papers [7, 8] and [26].

References

[1] I. D. Aron and S. Gupta, "Analytical comparison of local and end-to-end error recovery in reactive routing protocols for mobile ad hoc networks," in *Proc. ACM Workshop on Modeling, Analysis and Sim. of Wireless and Mobile Systems (MSWiM)*, (Boston, MA), 2000.

[2] C. Bettstetter, "Mobility modeling in wireless networks: Categorization, smooth movement, and border effects," *ACM Mobile Comp. and Comm. Review*, vol. 5, no. 3, 2001.

[3] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM Intern. Symp. on Mobile Ad Hoc Netw. and Comp. (MobiHoc)*, (Lausanne, Switzerland), June 2002.

[4] C. Bettstetter, "Topology properties of ad hoc networks with random waypoint mobility," accepted for *Proc. ACM Intern. Symp. on Mobile Ad Hoc Netw. and Comp. (MobiHoc)*, poster session, (Annapolis, MD), June 2003.

[5] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa, "Stochastic properties of the random waypoint mobility model," accepted for *ACM/Kluwer Wireless Networks*, to appear 2004.

[6] C. Bettstetter and O. Krause "On border effects in modeling and simulation of wireless ad hoc networks," in *Proc. IEEE Intern. Conf. on Mobile and Wireless Comm. Netw. (MWCN)*, (Recife, Brazil), 2001.

[7] C. Bettstetter and C. Wagner, "The spatial node distribution of the random waypoint model," in *Proc. 1st German Workshop on Mobile Ad Hoc Networks (WMAN)*, (Ulm, Germany), 2002.

[8] D. M. Blough, G. Resta, and P. Santi, "A statistical analysis of the long-run node spatial distribution in mobile ad hoc networks," in *Proc. ACM Intern. Workshop on Modeling, Analysis, and Sim. of Wireless and Mobile Systems (MSWiM)*, 2002.

[9] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proc. ACM Intern. Conf. on Mobile Computing and Networking (MobiCom)*, (Dallas, TX), Oct. 1998.

[10] T. Camp, J. Boleng, V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communication & Mobile Computing (WCMC)*, vol. 2, no. 5, pp. 483-502, Wiley, 2002.

[11] S. R. Das, C. E. Perkins, E. M. Royer, and M. K. Marina, “Performance comparison of two on-demand routing protocols for ad hoc networks,” *IEEE Personal Communications*, pp. 16–28, Feb. 2001.

[12] T. D. Dyer and R. V. Boppana, “A comparison of TCP performance over three routing protocols for mobile ad hoc networks,” in *Proc. ACM Intern. Symp. on Mobile Ad Hoc Netw. and Comp. (MobiHoc)*, (Long Beach, CA), 2001.

[13] Y. Fang, I. Chlamtac, and Y.-B. Lin, “Portable movement modeling for PCS networks,” *IEEE Trans. on Vehicular Technology*, July 2000.

[14] M. Grossglauser and D. Tse, “Mobility increases the capacity of ad hoc wireless networks,” in *Proc. IEEE Infocom*, (Anchorage, AK), April 2001.

[15] R. A. Guérin, “Channel occupancy time distribution in a cellular radio system,” *IEEE Trans. on Vehicular Technology*, vol. 36, Aug. 1987.

[16] P. Gupta and P. R. Kumar, “Critical power for asymptotic connectivity in wireless networks,” in *Stoch. Analysis, Control, Optim. and Appl.*, Birkhäuser, 1998.

[17] P. Gupta and P. R. Kumar, “The capacity of wireless networks.” in *IEEE Trans. on Inf. Theory*, vol. 46, no. 2, March 2000.

[18] G. Holland and N. H. Vaidya, “Analysis of TCP performance over mobile ad hoc networks,” in *Proc. ACM Intern. Conf. on Mobile Comp. and Netw. (MobiCom)*, (Seattle, WA), Aug. 1999.

[19] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang, “A group mobility model for ad hoc wireless networks,” in *Proc. ACM Intern. Workshop on Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWiM)*, (Seattle, WA), 1999.

[20] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark, “Scenario-based performance analysis of routing protocols for mobile ad hoc networks,” in *Proc. ACM Intern. Conf. on Mobile Comp. and Netw. (MobiCom)*, (Seattle, WA), Aug. 1999.

[21] D. B. Johnson, and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," *Mobile Computing*, Kluwer, pp. 153–181, 1996.

[22] D. Lam, D. C. Cox, and J. Widom, "Teletraffic modeling for personal communication services," *IEEE Communications*, Oct. 1997.

[23] Z. Lei and C. Rose, "Probability criterion based location tracking approach for mobility management of personal communication systems," in *Proc. IEEE Globecom*, (Phoenix, Arizona), pp. 977–981, 1997.

[24] A. Nasipuri, R. Castaneda, and S. Das, "Performance of multipath routing for on-demand protocols in mobile ad hoc networks," *ACM/Kluwer Mobile Networks and Applications*, vol. 6, pp. 339–349, 2001.

[25] "The network simulator - ns-2." http://www.isi.edu/nsnam/ns/, June 2002.

[26] G. Resta and P. Santi, "An analysis of the node spatial distribution of the random waypoint model for ad hoc networks," to appear in *Proc. ACM Workshop on Principles of Mobile Computing (POMC)*, (Toulouse, France), Oct. 2002.

[27] E. M. Royer and C. E. Perkins, "Multicast operation of the ad hoc on-demand distance vector routing protocol," in *Proc. ACM MobiCom*, (Seattle, WA), Aug 1999.

[28] E. M. Royer, P. M. Melliar-Smith, and L. E. Moser, "An analysis of the optimum node density for ad hoc mobile networks," in *Proc. IEEE Intern. Conf. on Communications (ICC)*, (Helsinki, Finland), June 2001.

[29] L. A. Santaló, *Integral Geometry and Geometric Probability*, Addison-Wesley, 1976.

[30] P. Santi, D.M. Blough, "The critical transmitting range for connectivity in sparse wireless ad hoc networks", *IEEE Trans. on Mobile Computing*, January-March 2003.

[31] J. Song and L. E. Miller, "Empirical analysis of the mobility factor for the random waypoint model," In *Proc. OPNETWORK*, (Washington, DC), Aug 2002.

[32] M. Spohn and J. J. Garcia-Luna-Aceves, “Neighborhood aware source routing,” in *Proc. ACM Symp. on Mobile Ad Hoc Netw. and Comp. (MobiHoc)*, (Long Beach, CA), pp. 11–21, Oct 2001.

[33] Y. Xu, J. Heidemann, and D. Estrin, “Geography-informed energy conservation for ad hoc routing,” in *Proc. ACM MobiCom*, (Rome, Italy), 2001.

[34] J. Yoon, M. Liu, and B. Noble, “Random waypoint considered harmful,” in *Proc. IEEE Infocom*, (San Francisco, CA), April 2003.

[35] X. Zeng, R. Bagrodia, and M. Gerla, “GloMoSim: a library for parallel simulation of large-scale wireless networks,” in *Proc. Workshop on Parallel and Distr. Simulations (PADS)*, (Banff, Canada), May 1998.

[36] M. M. Zonoozi and P. Dassanayake, “User mobility modeling and characterization of mobility patterns,” *IEEE Journal on Sel. Areas in Communications*, vol. 15, pp. 1239–1252, Sept. 1997.