A Statistical Analysis of the Long-Run Node Spatial Distribution in Mobile Ad Hoc Networks

Douglas M. Blough* Giovanni Resta† Paolo Santi†

CONTACT AUTHOR: **Giovanni Resta**, Istituto di Informatica e Telematica del CNR, Via G. Moruzzi 1, 56124, Pisa (Italy). Tel. +39 050 3152408, Fax. +39 050 3152333, e-mail: giovanni.resta@iit.cnr.it.

Abstract

In this paper, we analyze the node spatial distribution of mobile wireless ad hoc networks. Characterizing this distribution is of fundamental importance in the analysis of many relevant properties of mobile ad hoc networks, such as connectivity, average route length, and network capacity. In particular, we have investigated under what conditions the node spatial distribution resulting after a large number of mobility steps resembles the uniform distribution. This is motivated by the fact that the existing theoretical results concerning mobile ad hoc networks are based on this assumption. In order to test this hypothesis, we performed extensive simulations using two well-known mobility models: the random waypoint model, which resembles intentional movement, and a Brownian-like model, which resembles non-intentional movement. Our analysis has shown that in Brownian-like motion the uniformity assumption does hold, and that the intensity of the concentration of nodes in the center of the deployment region that occurs in the random waypoint model heavily depends on the choice of some mobility parameters. For extreme values of these parameters, the uniformity assumption is impaired.

1 Introduction

Wireless ad hoc networks have received increasing interest in the scientific community in recent years. However, due to their complex and unstructured nature, very few analytical results describing their fundamental properties have been derived. Among them, the most notable concern the connectivity and coverage problems for stationary networks, which have been analyzed in [11, 21, 24, 25]. A common assumption in these studies is that nodes are distributed in a given area according to a probability distribution; the value of the node transmitting range ensuring a connected network (or coverage of the deployment area) with high probability is then derived. Another important theoretical result for stationary ad hoc networks is presented in [12], where it is shown that the capacity of the network does not scale with its size.

If deriving analytical results for stationary networks is difficult, even more challenging is deriving theoretical results regarding *mobile* ad hoc networks. This has been done, for example, in [10], where it is shown that, contrary to the stationary case, the capacity of mobile ad hoc networks can actually scale with the size (if high communication latency can be tolerated). Less general analytical results for mobile networks have been derived in [1, 20], where two variations of the DSR routing protocol are evaluated in a theoretical framework.

As in the case of stationary networks, these results rely on some assumptions on the node spatial distribution, which are designed to simplify the analysis. In particular, in [1, 10] it is assumed that:

– (*uniformity assumption*) in any snapshot of the mobile network the nodes are distributed uniformly and independently at random in the deployment area;

– (*independence assumption*) different snapshots are independent.

*School of ECE, Georgia Institute of Technology, Atlanta, GA. e-mail: doug.blough@ece.gatech.edu

†Istituto di Informatica e Telematica del CNR, Pisa, Italy. e-mail: {giovanni.resta,paolo.santi}@iit.cnr.it

With these hypotheses, a mobile network composed by n nodes performing N mobility steps can be modeled as N independent experiments, where each experiment consists in distributing the n nodes uniformly in the deployment region.

Although these assumptions may be realistic in some scenario, their likelihood heavily depends on the initial node distribution and on the mobility pattern considered. It follows that the choice of the mobility model and/or of the initial node distribution could have a dramatic impact on the accuracy of the analytical results of [1, 10]. Hence, the influence of the mobility pattern and/or of the initial placement on the node spatial distribution that results after a large number of mobility steps must be precisely evaluated.

Probabilistic mobility modeling has been extensively studied in the related field of cellular networks. For example, Zonoozi and Dassanayake [28] have shown (by means of extensive simulations and "goodness-of-fit" tests) that the cell residence time of a mobile user moving according to a given mobility pattern follows a generalized gamma distribution. In the Brownian mobility model, it has been shown that, given the user's location at time t_0, the probability distribution of the physical location at time $t > t_0$ can be calculated [17]. For a survey of probabilistic mobility modeling in cellular networks see [3].

Although interesting, these results are not directly applicable to ad hoc networks. In fact, most of them concern specific properties of cellular networks (e.g., cell residence time, number of handoffs during a call, and so on). Furthermore, the emphasis is usually on modeling the properties of a single mobile user, rather than on the global mobile users' distribution. Finally, the mobility patterns considered resemble the typical motions that occur in cellular systems, i.e. human and vehicular motion. Hence, an accurate study of fundamental properties of a mobile ad hoc network is, to the best of our knowledge, still lacking.

Preliminary steps in this direction have been recently done in [3, 4], where it has been observed that in the well known random waypoint model [13] nodes tend to concentrate in the center of the deployment region, causing the so called *border effect*. However, the intensity of this border effect has not been carefully evaluated, and in particular its dependence on the choice of the mobility parameters has not been investigated.

In this paper, we perform a further step in understanding the node spatial distribution generated by mobile ad hoc networks. In particular, we study under what conditions is the uniformity assumption used to derive the analytical results of [1, 10] valid. The mobility patterns considered, which are described in detail in Section 3, account for two different kinds of motion: *intentional*, which is modeled using the random waypoint model, and *non-intentional*, which is modeled by a Brownian-like motion. In the case of the random waypoint model, we investigate the conditions (i.e., the choices of the mobility parameters) under which the spatial distribution (which *is not* uniform) can be well approximated by the uniform distribution.

In order to test the uniformity of the node spatial distribution, we have designed a collection of statistical tests tailored to wireless ad hoc networks. Each of these tests measures a feature of the node distribution that is of interest in the design of wireless ad hoc networks. Details on the uniformity test used in this paper can be found in Section 2.

The results of our analysis show that in Brownian-like motion the uniformity assumption does hold, while the intensity of the border effect in the random waypoint model heavily depends on the values of the mobility parameters. In particular, it becomes more and more intense as the pause time (i.e., the resting time between two consecutive movements) decreases. For extreme values of the pause time (e.g., zero), the uniformity assumption is impaired. The intensity of the border effect is also affected, to a lesser extent, by the fraction of stationary nodes (i.e., of the nodes that do not move during the entire simulation time), while it is practically independent of the velocity of nodes.

In addition to its intrinsic importance as a fundamental property of mobile ad hoc networks, the study of the node spatial distribution is significant in a number of ways. For example, the distribution of route lengths is of critical importance in determining the performance of routing algorithms in presence of mobility [1, 13, 20]. Longer routes have a higher probability of becoming invalid due to node mobility than shorter ones. When routes become invalid, these algorithms incur overhead in discovering new ones. Observe that the route length distribution can be determined directly from the node spatial distribution when all nodes have the same transmitting range.

Another area in which node spatial distribution is highly relevant is in the evaluation of network capacity, where route length and node spatial distribution are important factors [10, 12, 16]. For example, a node concentration in the center of the deployment region, as it is observed in the random waypoint model, reduces average route lengths, which tends to increase capacity. However, this concentration also exacerbates interference, which has the opposite effect of decreasing capacity. A detailed modeling and evaluation of capacity in presence of mobility is beyond the scope of this paper, but quantifying the node spatial distribution should make more accurate capacity evaluation possible in the future.

2 An uniformity test for ad hoc networks

Statistical tests are commonly used to verify the conformity of a large set of observations to a given distribution. A typical test is as follows. Suppose H is the statistical hypothesis that we want to test, i.e. that a sample consisting of a large number of independent observations $\mathcal{X} = \{x_1, \ldots, x_l\}$ is taken from the continuous distribution $F(x)$. Let μ be a statistic calculated on the sample, and assume that, in the hypothesis that observations are distributed according to $F(x)$, the limit distribution of μ, i.e. the probability distribution of μ when $l \to \infty$, is known. Let CDF_μ be the cumulative distribution function of this distribution, i.e. $CDF_\mu(x) = P(\mu \leq x)$, and let $\bar{\mu}$ be the value of the statistic calculated on $\mathcal{X}$. The test is as follows: if $C_1 \leq \bar{\mu} \leq C_2$ the hypothesis H is adopted, otherwise it is rejected. If l is sufficiently large, the error of the first kind, i.e., the probability of rejecting H when the hypothesis is true, is well approximated by $P(\{\mu < C_1\} \bigcup \{\mu > C_2\}) = 1 + CDF_\mu(C_1) - CDF_\mu(C_2)$. By appropriately choosing constants C_1 and C_2, we get the desired test accuracy.

Observe that the test described above provides only unidirectional information. In fact, if the sample is rejected, then we know that it is very unlikely that the hypothesis H actually holds. However, if the test passes, we can only conclude that the sample does not violate H and this, in general, is not sufficient to conclude that H actually holds.

To clarify this point, consider the following example. Suppose that observations are a set of points taken from the closed interval $[0, 1]$, and assume that the hypothesis H is that the points are uniformly and independently distributed in that interval. Let μ be the statistic that measures the distance between the extreme points in the sample, and let $E[\mu]$ be its expectation under the hypothesis H. Consider now a sample $\bar{\mathcal{X}}$ in which a point is located at coordinate 0, another is located at coordinate $E[\mu]$, and the remaining $l - 2$ points are distributed uniformly and independently in $(0, E[\mu])$. Clearly, the sample passes the test, but hypothesis H is violated. In order to mitigate this problem, more than one statistic can be considered: if the sample passes all the tests it is accepted, otherwise it is rejected. Returning to the example above, if we also calculate the sum of the values of the points, the sample $\bar{\mathcal{X}}$ will be probably rejected.

The discussion above indicates that the test accuracy can be significantly improved by considering different statistics on the same sample. In fact, while it is relatively likely that a sample which does not conform to the hypothesis H fools one test, it is much less likely that it passes several tests on different statistics. Hence, a good statistical test must be based on many statistics, which must be adequately chosen given the nature of the observations and the hypothesized distribution.

In this paper, an observation consists of the positions of the nodes of a mobile ad hoc network at an arbitrary instant of time t. We assume that the network is composed by n nodes, which move in a square region of unit area. Hence, the observation taken at time t, also called *snapshot* at time t, can be formally defined as the set $\mathcal{X}(t) = \{x_1(t), \ldots, x_n(t)\}$, where $x_i(t)$ are the coordinates of the point in $[0, 1]^2$ that represent the position of node i at time t.

The statistical hypothesis H that we want to test is that the points in $\mathcal{X}(t)$ are independently and uniformly distributed in the deployment region $[0, 1]^2$. To this purpose, we will consider the following statistics on $\mathcal{X}(t)$:

1. the length of the longest nearest neighbor edge, defined as

$$LNN_t = \max_{i=1,\ldots,n} \min_{j \neq i} \{d(x_i(t), x_j(t))\},$$

   where $d(x_i(t), x_j(t))$ denotes the Euclidean distance between points $x_i(t)$ and $x_j(t)$;

2. the length of the longest edge of the Euclidean minimum spanning tree (MST) built on the points in $\mathcal{X}(t)$;

3. the total edge length of the MST on $\mathcal{X}(t)$;
4. the total edge length of the Voronoi diagram on $\mathcal{X}(t)$;
5. the total edge length of the Delaunay triangulation on $\mathcal{X}(t)$;
6. the number of empty cells.

The Voronoi diagram and Delaunay triangulation are defined as follows [9].

Given a set $\mathcal{X}$ of points in the plane, the Voronoi diagram partitions the plane into regions, one per point; the region for a point x consists of all points closer to x than to any other point. The boundary of every region is a (possibly infinite) polygon, whose sides are called *Voronoi edges*. In our test we will measure the total length of the internal Voronoi edges, i.e. of those edges that do not intersect the boundary of the deployment region $[0, 1]^2$.

The Delaunay triangulation, which is the dual of the Voronoi diagram, is the unique triangulation such that the circumcircle of every triangle contains no points of $\mathcal{X}$ in its interior. It is known that the MST on $\mathcal{X}$ is a subgraph of the Delaunay triangulation [9]. The Voronoi Diagram, Delaunay triangulation, and MST of a set of points are shown in Figure 1.

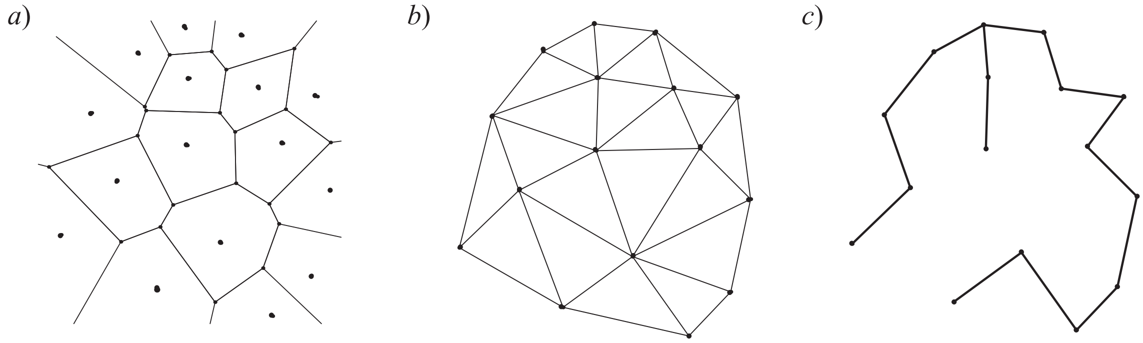

Figure 1: Voronoi diagram (a), Delaunay triangulation (b) and MST (c) of a set of points in the plane.

The empty cell statistic is obtained by partitioning the deployment region into n square cells of side $\frac{1}{\sqrt{n}}$ that are arranged in a grid fashion, and by counting the number of vacant cells.

The statistics above have been chosen since their limit distributions are known, and since they are relevant parameters in wireless ad hoc networks. In fact, the length of the longest nearest neighbor edge is a lower bound to the critical transmitting range[1], while the length of the longest edge in the MST is exactly the critical transmitting range [7, 22]. The critical transmitting range in ad hoc networks has been studied in [11, 21, 25]. The total edge length of the MST is closely related to the cost of the optimal transmitting range assignment [5, 14], while the Voronoi diagram and Delaunay triangulation, or derivation of these graphs, are used to construct spanners for power-efficient routing [8], and in the evaluation of sensor networks coverage and exposure [18, 19]. Finally, a subdivision of the deployment region into square cells is used, for instance, to implement cooperative strategies aimed at extending network lifetime [6, 27].

The limit distributions of these statistics, under the assumption that nodes are distributed uniformly and independently at random in $[0, 1]^2$, have been widely studied. In particular, it has been shown that:

- the longest nearest neighbor edge and the longest edge of the MST have the same limit distribution, which is double exponential. Denoting with NNL and LMST the random variables corresponding to these statistics, we have [7, 22]:

$$\lim_{n\to\infty} P(NNL \leq \sqrt{\frac{\alpha + \log n}{\pi n}}) = \lim_{n\to\infty} P(LMST \leq \sqrt{\frac{\alpha + \log n}{\pi n}}) = \exp(-e^{-\alpha})$$

[1]The critical transmitting range is the minimum value r of the transmitting range that ensures that the communication graph obtained by connecting nodes at distance at most r is connected.

- denoting by $TMST$ the total edge length of the MST, we have [2, 26]:

$$\lim_{n\to\infty} \frac{TMST}{\sqrt{n}} = c,$$

where c is a constant, with $0.600822 \le c \le 0.707$. The value of c has been experimentally evaluated as ≈ 0.656. Although a formal proof does not exist, there are many evidences that the limit distribution of $TMST$ is the normal distribution [26].

- the limit distribution of the total edge length of the Voronoi diagram and of the Delaunay triangulation is the normal distribution [23].

- the limit distribution of the number of empty cells is the normal distribution of parameters $(\frac{n}{e}, \sqrt{\frac{n(1-(2/e))}{e}})$ [15].

In the following sections, we will use these statistics to test our hypothesis H.

3 The mobility models

Two mobility models will be used in the following to test the statistical hypothesis described in the previous Section.

The first model is the random waypoint model [13], which is commonly used to evaluate the performance of protocols for ad hoc networks in presence of mobility. In this model, every node chooses uniformly at random a destination in $[0,1]^2$, and moves towards it along a straight line with a velocity chosen uniformly at random in the interval $[v_{min}, v_{max}]$. When it reaches the destination, it remains stationary for a predefined pause time t_{pause}, and then it starts moving again according to the same rule. We have also included a further parameter in the model, namely the probability p_{stat} that a node remains stationary during the entire simulation time. Hence, only $(1-p_{stat}) \cdot n$ nodes (on the average) will move. Introducing p_{stat} in the model accounts for those situations in which some nodes are not able to move. For example, this could be the case when 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 where two types of nodes are used, one type that is stationary and another type that is mobile.

It has been observed in [3] that nodes moving according to the random waypoint model tend to concentrate in the center of the deployment region. This phenomenon, called the *border effect*, is due to the fact that when the starting and ending point of the node's movement are chosen uniformly at random in a bounded region, it is very likely that the trajectory that connects them crosses the center of the region. This can be better understood considering the angle of the trajectory at the node's starting point: while nodes in the middle of the area have uniformly distributed angle, nodes near the border are more likely to choose an angle that leads them back in the center of the region.

Depending on the intensity of the border effect, the node spatial distribution generated by a mobile network could be rejected by at least one of the statistical tests for uniformity described above. As we will see in Section 5, the intensity of the border effect is closely related to some parameters of the random waypoint model and, for extreme values of these parameters, it causes the failure of the uniformity test for at least one statistic.

The second mobility model considered in this paper resembles the Brownian two-dimensional motion. Mobility is modeled using parameters p_{stat}, p_{move} and m. Parameter p_{stat} is defined as in the random waypoint model. Parameter p_{move} is the probability that a node moves at a given step. This parameter accounts for heterogeneous mobility patterns, in which nodes may move at different times. Intuitively, the smaller is the value of p_{move}, the more heterogeneous is the mobility pattern. However, values of p_{move} close to 0 result in an almost stationary network. If a node is moving at step i, its position in step $i+1$ is chosen uniformly at random in the square of side $2m$ centered at the current node location. Parameter m models, to a certain extent, the velocity of the nodes: the larger m is, the more likely it is that a node moves far away from its position in the previous step.

Observe that, contrary to the case of the random waypoint model, in the Brownian-like motion it is possible that a moving node close to the border chooses as its next position a point which is outside the boundary of the deployment region. To circumvent this problem, we need a so called *border rule* [3], which defines what to do with these nodes. In this situation, the node can be:

- bounced back according to some rule;
- positioned at the point of intersection of the boundary with the line connecting the current and the desired next position;
- wrapped around to the other side of the region, which is considered as a torus;
- "deleted", and a new node is initialized according to the initial distribution;
- forced to choose another position, until the chosen position is inside the boundaries of the deployment region.

Depending on the choice of the border rule, non-uniformity can be produced in the Brownian-like model also. For example, the second rule described above places nodes exactly on the boundary of the region with higher probability than at other points. This is the reverse of the border effect discussed so far, in which nodes are more likely to be near the center of the region. In fact, the only two rules that do not appear to favor one part of the region over another are the torus rule and the one in which a node is eliminated when it would cross the boundary and a new node is created in its place. The latter rule does not have any basis in the real world, so we do not consider it here. The torus rule would seem to apply to the real world only in some very specific applications and so we do not consider it either. There does not appear to be a large difference between the other models and so we arbitrarily chose the last rule to implement in our simulator. As we will see in Section 5, the intensity of the border effect in the Brownian-like model when using this border rule is very weak, and does not prevent the resulting node spatial distribution from passing all the uniformity tests.

4 The simulation model

To test the uniformity of the node spatial distribution of mobile ad hoc networks we have developed an ad hoc simulator. The simulator takes as input the number n of nodes to distribute, the number $\sharp sim$ of simulations to run, the number $\sharp steps$ of mobility steps to perform for each simulation, and the mobility model to use, along with its parameters. Nodes are distributed uniformly and independently at random in $[0,1]^2$; then, they start moving according to the specified mobility model. After $\sharp steps$ of mobility, the statistics used for the uniformity test (i.e., the length of the longest nearest neighbor and MST edge, the total edge length of the MST, Voronoi diagram and Delaunay triangulation, and the number of empty cells) are calculated. In order to calculate the empty cell statistic, the deployment region is divided into n square cells of side $\frac{1}{\sqrt{n}}$, and the number of vacant cells is counted. For the sake of simplicity, in our experiments we have always used perfect-square values for n, namely 49, 100 and 900.

The simulator returns the average value and the standard deviation over the $\sharp sim$ runs for each of the six statistics considered. These values are used to test the "uniformity" of the node spatial distribution as described in Section 2. The simulator also returns six output files containing all the $\sharp sim$ values of each statistic. These files are used for the single test evaluation described in Section 5.1. Finally, the simulator returns the occupancy file needed for the occupancy-based analysis described in Section 5.3.

5 Simulation results

In this section, we report the results of the extensive simulations we have performed. We ran three sets of simulations, the results of which are presented in separate subsections. The first set was aimed at determining the parameters of the limit distribution for some of the statistics considered, and at evaluating the accuracy of each single test. In the second set of simulations, we performed the uniformity test for networks of different sizes and using different mobility models and parameters. In the last subsection, we present the results of the occupancy-based analysis, aimed at carefully investigating the border effect.

5.1 Evaluating the test accuracy

Before testing the uniformity of mobile networks, we have performed preliminary simulations on stationary networks (which are obtained by setting $\sharp steps$ to 0).

The purpose of these simulations was to determine the parameters (expected value and standard deviation) of the limit distribution for the total edge length of the MST, Voronoi diagram, and Delaunay triangulation. We recall that, contrary to the case of the other statistics, the limit distribution of the statistics above is known to be normal, but the actual parameters are unknown. In order to be accurate, we considered a very large sample of 100000 simulations. The average value and standard deviation of these statistics for networks of size n=49, n=100 and n=900 have been determined, and are reported in Table 1. These values are used to derive the upper and lower bounds for the corresponding statistical test as described in Section 2.

n	Avg[MST]	Std[MST]	Avg[Vor]	Std[Vor]	Avg[Del]	Std [Del]
49	4.769	0.229	9.578	0.538	23.958	1.372
100	6.754	0.216	15.496	0.482	36.054	1.414
900	19.757	0.199	55.370	0.339	110.674	1.375

Table 1: Average value and standard deviation of the MST, Voronoi diagram, and Delaunay triangulation total edge length for uniformly distributed nodes.

Another goal of this set of preliminary simulations was to evaluate the rate of convergence of the experimental distribution obtained from a large sample of uniformly distributed points to the limit distribution. In fact, if the rate of convergence of the actual distribution to the limit distribution is very low, the uniformity test could be inaccurate. In other words, it could be the case that, for moderate values of n, the test (which is based on the limit distribution) is not passed even by a sample of uniformly distributed nodes.

The accuracy of the statistical tests was tested using the output files produced by the simulator for each single statistic. In order to keep the simulation time reasonable and the size of the output files manageable, we considered a network of n=100 nodes and a sample of 10000 experiments. From the data contained in each file, we generated a histogram representing the experimental density of the statistic. For the test to be accurate, this histogram should closely resemble the theoretical density function, which is obtained by the limit distribution[2].

As shown in Figure 2, this is actually the case for the MST, Voronoi diagram and Delaunay triangulation total edge length statistic, and for the empty cell statistic, meaning that the tests based on these parameters are very accurate. On the contrary, the longest nearest neighbor edge and, more notably, the longest MST edge statistic show significant discrepancy between the experimental and the theoretical distribution. As observed above, this is due to the fact that the rate of convergence to the limit distribution is very low. In the case of the longest nearest neighbor edge statistic, the resemblance is anyway sufficient to ensure the accuracy of the test: when the test is executed on small instances (n=49) of uniformly distributed nodes, it is passed. This is not the case of the longest MST edge statistic, for which the test on uniformly distributed nodes is not passed even for very large instances (n=900). Hence, the test based on this statistic is inaccurate, and will not be considered in the remainder of this paper.

5.2 The uniformity test

In the second set of simulations, we have performed the uniformity test on networks of size n=49, n=100 and n=900 using the two mobility models with different parameters.

First, we considered the random waypoint model, setting the mobility parameters as follows: t_{pause}=0, v_{min}=0.001, v_{max}=0.01, p_{stat}=0. Then, we varied each of these parameters separately (for the velocity,

[2]In the case of the MST, Voronoi diagram, and Delaunay triangulation total edge length statistic, we have used the parameters reported in Table 1.

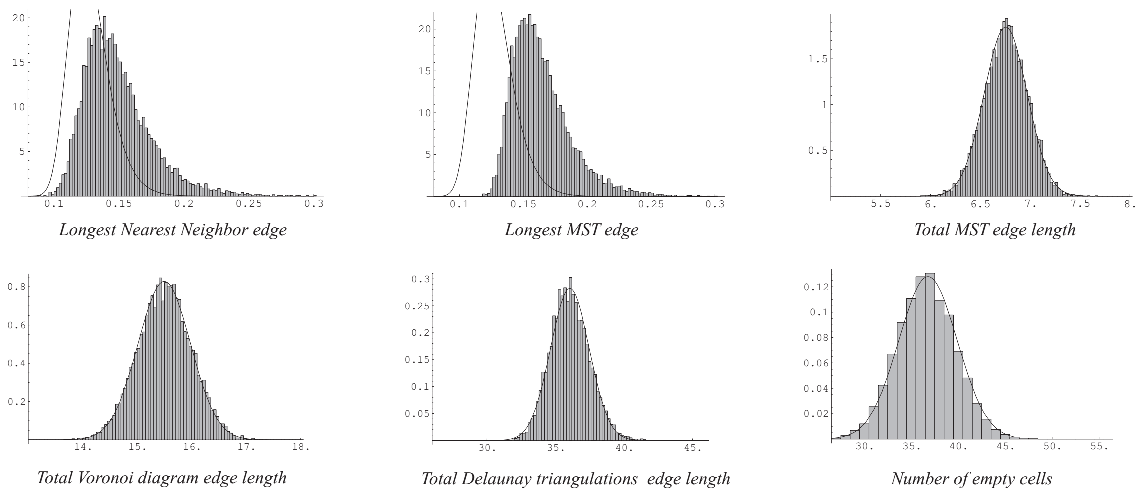

Figure 2: Experimental vs. theoretical density of the six statistics considered.

we considered only different values of v_{max})[3]. For each experiment, we ran 10000 simulations with 1000 steps of mobility each, and we performed all the statistical tests. The tests, all with a 95% accuracy, were performed on the average value (over the 10000 simulations) of the statistic of interest, measured on the node distribution resulting at the end of the 1000 mobility steps. These values for the number of mobility steps and sample size were chosen because they are a reasonable compromise between statistical accuracy and runtime.

The results of these experiments, which are reported in Table 2, have shown a significant effect of the pause time on the test outcomes: a longer pause time results in a more uniform node spatial distribution. A pause time of 200 steps (over 1000 steps of mobility) is sufficient to render the node spatial distribution indistinguishable from uniform, except for the longest nearest neighbor edge statistic when n=900. A similar, although less evident, effect on the node spatial distribution is displayed by p_{stat}. Conversely, the uniformity of the node spatial distribution appears to be completely independent of v_{max}. Indeed, this is due to the fact that in our simulations we varied the velocity when t_{pause} was set to 0, i.e., when the nodes move continuously during the entire simulation time. In this scenario, varying the velocity has no effect on the intensity of the border effect, which is responsible for the non-uniformity of the distribution. However, v_{max} has some influence on the uniformity when t_{pause} is greater than 0: with an increased velocity, nodes arrive quickly at destination, where they rest for t_{pause} steps; hence, the fraction of moving nodes in the network (which is responsible for the border effect) is reduced. The results of the further simulations that we have performed setting t_{pause} to 200 and varying v_{max}, which are reported in Table 3, have shown that the influence of the velocity on the node spatial distribution is however very limited: only when $n = 900$, an increased velocity induces a more uniform distribution (the LongestNN test is passed only when v_{max} is 0.05 or higher).

Regarding the size of the network, small networks (n=49) passed more uniformity tests than larger ones (n=100 and n=900). As discussed in Section 2, this does not necessarily imply that the node spatial distribution for small networks is "more uniform" than for large ones. It could also be the case that since the accuracy of the statistical test increases with n, (we recall that the test is based on the limit distribution), it succeeds in rejecting an "almost uniform" sample only for large values of n. In order to gain further insights on this point, we have performed the occupancy-based analysis, the results of which are presented in the next subsection.

A similar set of simulations was also performed using the Brownian-like mobility model. The default values of the mobility parameters were set to $p_{move} = 0.7$, $p_{stat} = 0$ and m=0.1. As in the case of the random waypoint simulations, we varied each parameter separately and we performed the five statistical tests. In particular, we varied p_{move} from 0.7 to 1 in steps of 0.1, p_{stat} from 0 to 0.5 in steps of 0.25, and m from 0.01 to 0.1 in steps of 0.03. The experiments produced the same positive test outcomes, i.e.,

[3]The unit of time is a "simulation step", and velocity is expressed as normalized spatial unit over simulation step.

		LongestNN			TotalMST			TotalVor			TotalDel			EmptyCells		
		49	100	900	49	100	900	49	100	900	49	100	900	49	100	900
t_{pause}	0	yes	no	no	no	no	no	yes	yes	no	no	no	no	yes	yes	no
	100	yes	yes	no	yes	yes	no	yes	yes	no	yes	yes	no	yes	yes	no
	200	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
	300	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
v_{max}	0.01	yes	no	no	no	no	no	yes	yes	no	no	no	no	yes	yes	no
	0.05	yes	no	no	no	no	no	yes	yes	no	no	no	no	yes	yes	no
	0.1	yes	no	no	no	no	no	yes	yes	no	no	no	no	yes	yes	no
p_{stat}	0	yes	no	no	no	no	no	yes	yes	no	no	no	no	yes	yes	no
	0.3	yes	no	no	yes	no	no	yes	yes	no	yes	yes	no	yes	yes	no
	0.5	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	yes	yes

Table 2: Outcomes of the statistical tests on networks of size n=49, 100 and 900 with different mobility parameters.

		LongestNN			TotalMST			TotalVor			TotalDel			EmptyCells		
		49	100	900	49	100	900	49	100	900	49	100	900	49	100	900
v_{max}	0.01	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
	0.05	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
	0.1	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Table 3: Outcomes of the statistical tests on networks of size n=49, 100 and 900 for increasing values of v_{max}. The values of the other mobility parameters are $t_{pause} = 200$, $v_{min} = 0.001$ and $p_{stat} = 0$.

all the statistical tests were passed for each combination of the mobility values considered. These results confirm the intuition that Brownian-like motion has a "more uniform" nature compared to the random waypoint model.

The results of the statistical tests presented above give only a qualitative analysis of the node spatial distribution of mobile ad hoc networks. In fact, they show the conditions under which the statistics of interest measured on a set of nodes distributed according to a mobile distribution are consistent with those that would be obtained if the nodes' distribution was uniform. However, they give no numerical estimate of how close they actually are. In order to better evaluate this aspect, we performed a further set of simulations. We considered each statistic separately, and we compared the theoretical probability density function (obtained by the limit distribution) with the experimental density obtained in different mobile network scenarios. Since the resemblance of the experimental distribution with the uniform distribution in the random waypoint model is not always guaranteed, we focused our attention on this mobility model. In particular, we simulated two mobile scenarios for a network with n=100 nodes: one which is critical from the uniformity's point of view (t_{pause}=0, v_{min}=0.001, v_{max}=0.01 and p_{stat}=0), and the other for which all the uniformity tests are passed (t_{pause}=200, v_{min}=0.001, v_{max}=0.01, p_{stat}=0.3). As in the previous simulations, the size of the sample was 10000, and the statistics were calculated at the end of the 1000 mobility steps.

The results of these simulations are reported in Figure 3. Each graphic reports three curves: the

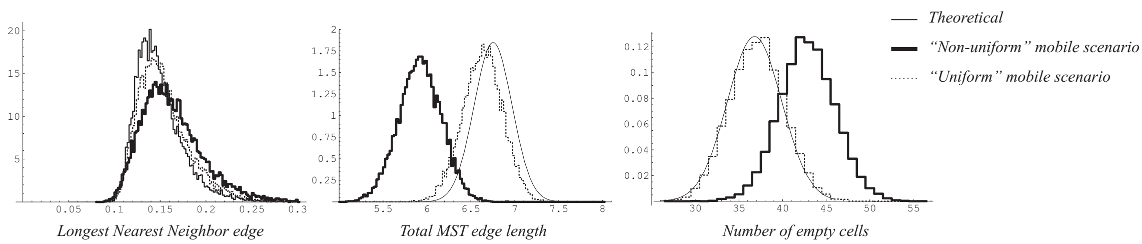

Figure 3: Density functions of the longest nearest neighbor edge, total MST edge length, and number of empty cells for the theoretical distribution, for two different mobile scenarios.

n	Avg[MST]	Std[MST]	Avg[Vor]	Std[Vor]	Avg[Del]	Std [Del]	Avg[Empty]	Std[Empty]
100	5.919	0.245	14.929	0.547	31.338	1.497	42.888	3.207

Table 4: Average value and standard deviation of the MST, Voronoi diagram, and Delaunay triangulation total edge length, and number of empty cells for the "non-uniform" mobile scenario.

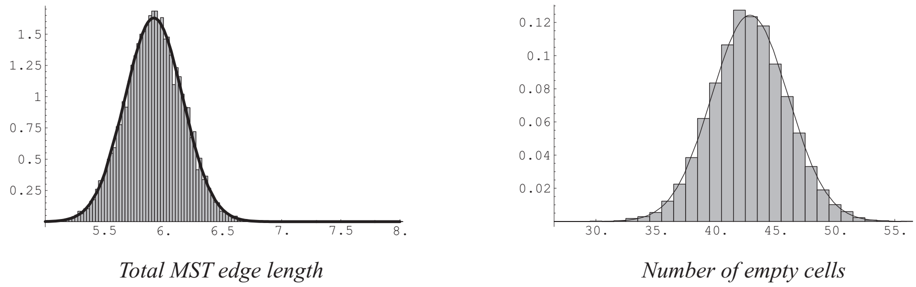

Figure 4: Experimental and normal distributions for the TotalMST and EmptyCell statistic.

theoretical density, and the experimental density in the "non-uniform" and "uniform" mobile scenarios. Due to the low rate of convergence to the limit distribution, for the LongestNN statistic we have reported, instead of the theoretical density, the experimental density obtained from a very large sample of uniformly distributed points. The graphics for the total Voronoi diagram and Delaunay triangulation edge length are not shown, since they are very similar to that of the total MST edge length.

As it is seen, the density in the mobile "non-uniform" scenario is similar in shape to the theoretical density, but it is shifted and somewhat scaled with respect to the theoretical curve. This seems to indicate that, also in presence of a relevant border effect, the limit distribution of the statistics considered is the same (normal for the total MST, Voronoi diagram and Delaunay triangulation edge length, and for the number of empty cells; double exponential for the longest nearest neighbor edge). However, the parameters of these limit distributions (expected value and standard deviation) are different. This fact, at least for the total MST, Voronoi diagram, and Delaunay triangulation edge length, and for the empty cell statistic, is confirmed by the very close resemblance of the experimental distribution to the normal distribution, where the parameters of the latter distribution are the average value and the standard deviation obtained by the experimental data (which are reported in Table 4). The experimental distributions for the TotalMST and EmptyCell statistics, and the corresponding normal distribution with the parameters set as in Table 4, are reported in Figure 4. An interesting open issue is characterizing the parameters of the distribution as a function of the mobility parameters.

Returning to Figure 3, we observe that the expected number of empty cells with "non-uniform" mobility is larger than in the uniform case (this easily follows from the fact that nodes are concentrated in the center of the deployment region), while the expected total MST edge length (as well as the total Voronoi diagram and Delaunay triangulation edge length) is smaller than in the uniform case. This fact merits some discussion. The nodes' concentration in the center of the deployment region caused by the border effect tends, on one hand, to reduce the length of the "inner" edges of the graph[4]. On the other hand, the "outer" edges, i.e., those close to the boundary, tend to be longer. The results of our simulations have shown that, on the average, the overall length reduction of "inner" edges outweighs the increment of the "outer" edges. A final observation regards the Voronoi diagram statistic, which is calculated considering only the *internal* Voronoi edges, i.e., those edges that do not intersect the boundary of the deployment region. In this case, the difference between the theoretical and the "non-uniform" mobile density is significantly reduced.

With "uniform" mobility, the experimental density was always very similar to the theoretical (or, in the case of the LongestNN statistic, to the experimental uniform) density. In the case of the Voronoi diagram statistic, the two densities were almost identical.

[4]With the word graph, we intend here either the MST, the Voronoi diagram or the Delaunay triangulation.

Altogether, the results presented in this subsection have shown that, if the border effect is not too intense, the node spatial distribution of a mobile network can be well approximated by the uniform distribution. If the border effect is very intense, the similarity with the uniform distribution is no longer valid. However, the limit distributions of many statistics of interest calculated on these networks remain the same, although with different parameters.

5.3 Occupancy-based analysis

The results presented in Section 5.2 have proved that in the Brownian-like model and, for appropriate values of the mobility parameters, in the random waypoint model, the node spatial distribution resulting after a large number of mobility steps is compatible with the uniform distribution. Once more, we want to emphasize that this does not mean that the distribution *is* uniform, but that, from the point of view of the five statistics considered, it is indistinguishable from the uniform distribution. On the other hand, it is known that, at least in the random waypoint model, nodes are subject to the border effect, which tends to concentrate them in the center of the deployment region. In order to better evaluate and "visualize" this phenomenon, we have performed a last set of simulations.

As in the case of the empty cell test, we subdivided the deployment region into n square cells of side $\frac{1}{\sqrt{n}}$, that are arranged in a grid fashion. Once the nodes are distributed uniformly at random in the region, they move according to the chosen mobility model for 1000 steps. At the end of the mobility steps, the number of nodes in each cell is recorded. Each experiment consisted of 10000 such simulations, and the simulator reported the cumulative number of nodes in each cell. If the node spatial distribution was uniform, this number should be approximately 10000 in every cell (we have n nodes to be distributed in n cells, hence the expected number of nodes in a cell for each simulation is 1). However, if the border effect occurs, cells in the center of the region should record a larger number of nodes than boundary cells.

The outcomes of these simulations have been used to build an "occupancy graphic", i.e. a three-dimensional graphic that for each cell in the 2D-plane, reports the cumulative number of nodes as its value in the third dimension. These graphics allow us to "visualize" the border effect.

As in the case of the previous simulations, we considered the two mobility models, and we assigned default values (the same as in Section 5.2) to the mobility parameters. Then, we varied each parameter separately, and we built the occupancy graphic.

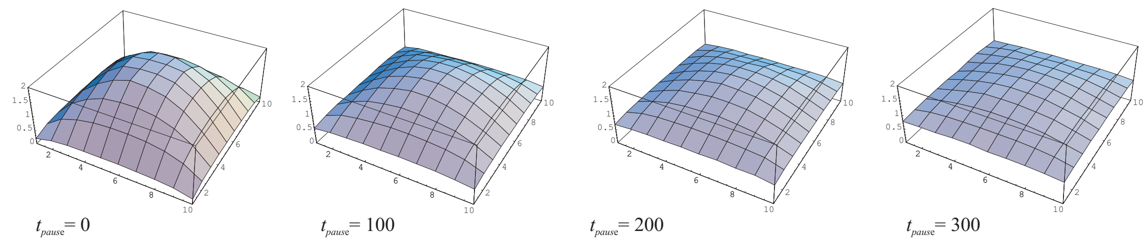

Figure 5: Occupancy graphics for the random waypoint model with different pause times.

The occupancy graphics obtained for the random waypoint model with pause times ranging from 0 to 300 are reported in Figure 5 (note that the values on the z-axis are normalized). The size of the network was n=100. The border effect is evident when $t_{pause} = 0$: the average number of nodes (over the 10000 simulations) in the corner cells is approximately 0.19, while in central cells this number is above 2, i.e. it is more than ten times larger than in the corner cells. As the pause time increases, the node spatial distribution becomes more and more flat: when $t_{pause} = 300$, the average number of nodes in corner cells is approximately 0.76, while in the central cells it is approximately 1.31, and the ratio between these

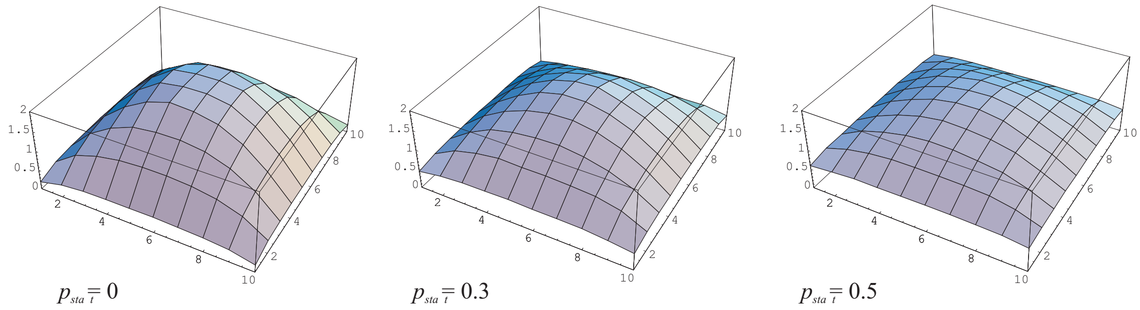

Figure 6: Occupancy graphics for the random waypoint model with different values of p_{stat}.

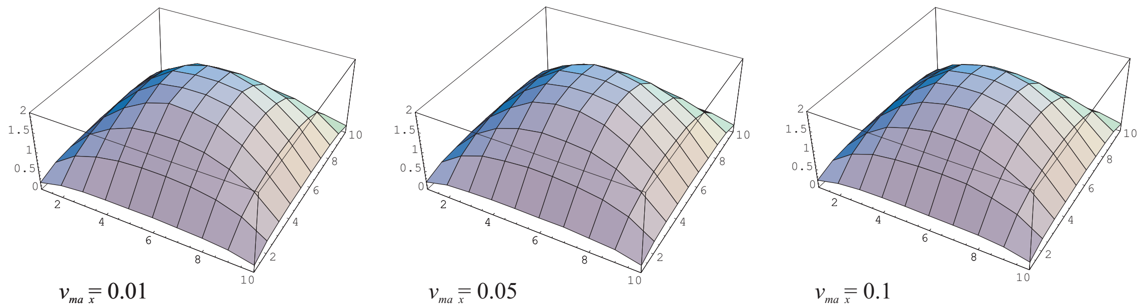

Figure 7: Occupancy graphics for the random waypoint model with different values of v_{max}.

numbers is below 2. A similar behavior is displayed for increasing values of p_{stat} (see Figure 6), while the effect of v_{max} on the node spatial distribution was negligible (see Figure 7). As discussed in Section 5.2, this is due to the fact that t_{pause} was set to 0. In case the pause time is not zero, an higher value of v_{max} makes the distribution more uniform (see also Table 3). Observe that for many combinations of the mobility parameters, the node distribution generated by the random waypoint model is very close to uniform. This is evident from the occupancy graphic of Figure 8, where parameters are set as in the "uniform" mobility scenario described in Section 5.2 (t_{pause}=200, v_{min}=0.001, v_{max}=0.01, p_{stat}=0.3).

In order to compare the effect of the different mobility parameters on the node spatial distribution quantitatively, we have also plotted the central cross-section of each of the occupancy graphics. The curves obtained are reported in Figure 9, and have confirmed the substantial influence of the pause time and, to a lesser extent, of p_{stat} on the node spatial distribution; conversely, v_{max} has no effect on the intensity of the border effect.

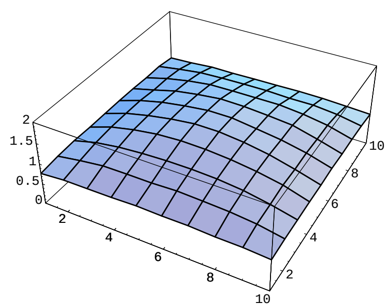

Figure 8: Occupancy graphics for the random waypoint model with parameters set as follows: t_{pause}=200, v_{min}=0.001, v_{max}=0.01, p_{stat}=0.3.

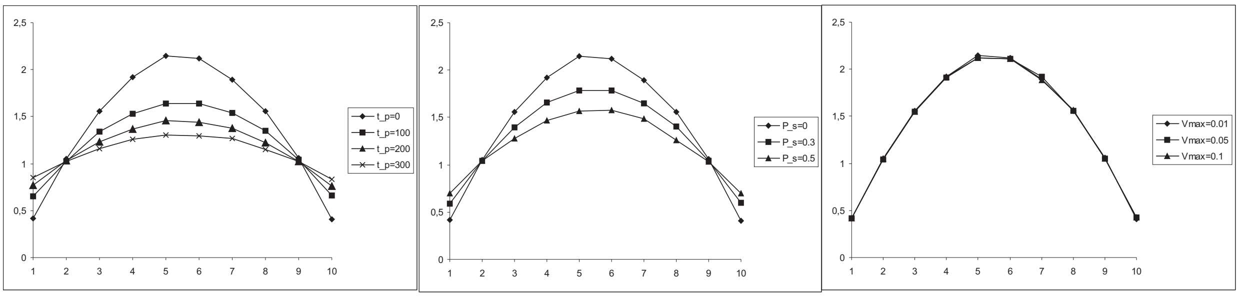

Figure 9: Central cross-section of the occupancy graphics for the random waypoint model with different values of the mobility parameters.

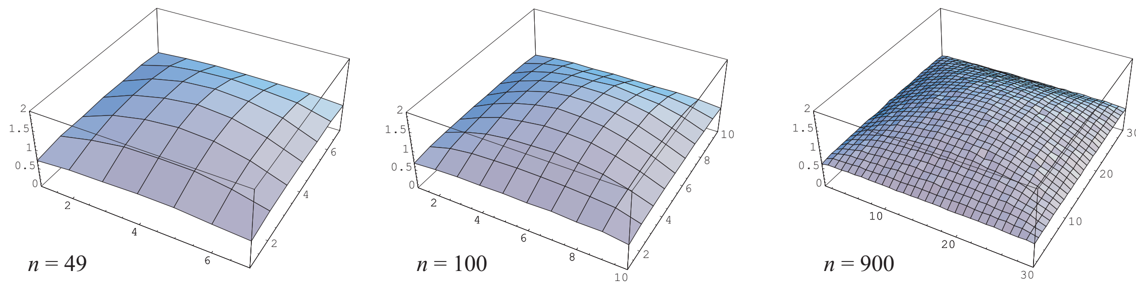

Figure 10: Occupancy graphics for the random waypoint model for networks of different size.

We have also investigated whether the intensity of the border effect is influenced by the size of the network or not. To this purpose, we set t_{pause}=200, $p_{stat} = 0$, $v_{min} = 0.001$ and $v_{max} = 0.01$, and we simulated networks of sizes n=49, n=100 and n=900. The occupancy graphics obtained by these simulations are reported in Figure 10. As it is seen, the intensity of the border effect seems to be virtually independent of the size of the network. However, a slight increment of the intensity for large networks is observed: while the ratio of the maximally over the minimally occupied cell is 2.027 when n=49, it is 2.191 when n=100, and 2.399 when n=900.

Finally, we simulated the Brownian-like mobility model. We considered networks of size n=100, we set the default values of the mobility parameters as in Section 5.2, and we varied each parameter separately. As expected, the node spatial distribution was practically indistinguishable from the uniform distribution for many combinations of the mobility parameters. The only parameter that displayed some influence on the node spatial distribution is m, whose increment causes a moderate border effect. This is due to the fact that, as m increases, nodes more and more distant from the border may choose as next position a point which is outside the boundaries of the deployment region. Hence, the "biasing towards the center" caused by the border rule becomes more and more intense. The occupancy graphics generated for different values of m are reported in Figure 11. As it is seen, the border effect is perceivable for m=0.05 and, more clearly, for m=0.1. However, it seems to be very different in nature from the border effect displayed by networks whose nodes move according to the random waypoint model. In fact, with random waypoint mobility nodes tend to be "smoothly" concentrated in the center of the deployment region; i.e., the node density is low at the borders, it becomes larger and larger as the cell is closer to the center, and is maximum in the central cells. What is observed here, instead, is a "sharp" and "localized" border effect: the node density is low in boundaries cells, but it is practically uniform in the rest of the deployment area.

As a final comment, it should be observed that, while the "smooth" border effect generated by the random waypoint model renders the node spatial distribution remarkably different from the uniform distribution (we recall that when t_{pause}=0 many uniformity tests are not passed), this is not the case of the "sharp" border effect generated by the Brownian-like motion, for which the resulting node spatial distribution is always indistinguishable from uniform.

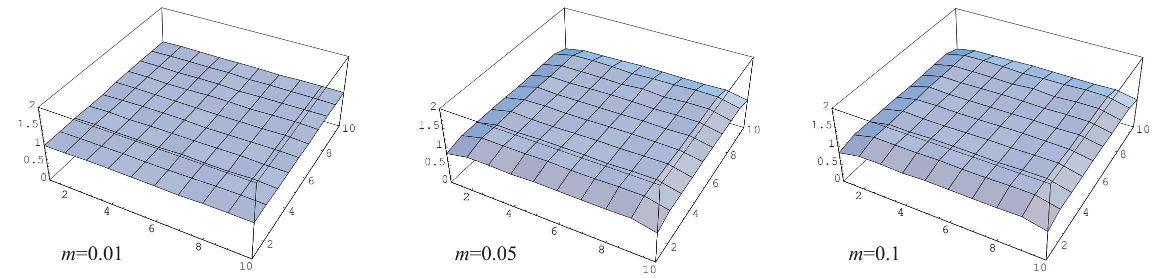

Figure 11: Occupancy graphics for the Brownian-like mobility model with different values of m.

6 Conclusions

In this paper we have investigated the node spatial distribution generated by a wireless ad hoc network whose nodes move for a large number of steps. In particular, we have tested whether this distribution can be approximated by the uniform distribution or not.

The results of the extensive simulations presented in this paper have shown that the truth of this hypothesis depends on the mobility model adopted: if we consider Brownian-like motion, the hypothesis holds, except for a slight and "sharp" border effect; conversely, in the random waypoint model the hypothesis holds only when the border effect is not too intense. The intensity of the border effect depends on the pause time and, to a lesser extent, on p_{stat}, while it is only marginally influenced by the velocity of the nodes. For most typical settings of the mobility parameter (e.g., $t_{pause} = 200$), the uniformity assumption holds, at least as far as the five statistics considered in this paper are considered. As a side effect, the results presented in this paper have indicated that the limit distribution of these statistics are the same as those resulting from a uniform node distribution, except for the values of their parameters (expected value and standard deviation).

We believe that the results presented in this paper are a further step in understanding the fundamental properties of a mobile ad hoc network. As further directions of research that this work has originated, we mention considering the impact of different initial node distributions (besides the uniform distribution considered in this paper) on the final spatial distribution, and considering group mobility models. Another very interesting direction of research is trying to formally characterize the spatial distribution generated by the mobility models considered in this paper.

References

[1] I.D. Aron, S. Gupta, "Analytical Comparison of Local and End-to-End Error Recovery in Reactive Routing Protocols for Mobile Ad Hoc Networks", *Proc. ACM Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 69–76, 2000.

[2] F. Avram, D. Bertsimas, "The Minimum Spanning Tree Constant in Geometrical Probability and Under the Independent Model: a Unified Approach", *The Annals of Applied Probability*, Vol. 2, n. 1, pp. 113–130, 1992.

[3] C. Bettstetter, "Smooth is Better than Sharp: A Random Mobility Model for Simulation of Wireless Networks", *Proc. ACM Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 19–27, 2001.

[4] C. Bettstetter, O. Krause "On Border Effects in Modeling and Simulation of Wireless Ad Hoc Networks", *Proc. 3rd IEEE International Conference on Mobile and Wireless Communication Networks (MWCMN)*, 2001.

[5] D.M. Blough, M. Leoncini, G. Resta, P. Santi, "On the Symmetric Range Assignment Problem in Wireless Ad Hoc Networks", *to appear in Proc. 2nd IFIP International Conference on Theoretical Computer Science*, Montreal, August 2002.

[6] D.M. Blough, P. Santi, "Investigating Upper Bounds on Network Lifetime Extension for Cell-Based Energy Conservation Techniques in Stationary Ad Hoc Networks", *Proc. ACM Mobicom 2002*, Atlanta, pp. 183–192, Sept. 2002.

[7] H. Dette, N. Heinze, "The Limit Distribution of the Largest Nearest-Neighor Link in the Unit d-Cube", *Journal of Applied Probability*, Vol. 26, pp. 67–80, 1989.

[8] J. Gao, L. Guibas, J. Hershberger, L. Zhang, A. Zhu, "Geometric Spanner for Routing in Mobile Networks", *Proc. ACM MobiHoc01*, pp. 45–55, 2001.

[9] J.E. Goodman, J. O'Rourke, *Handbook of Discrete and Computational Geometry*, CRC Press, New York, 1997.

[10] M. Grossglauser, D. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks", *Proc. IEEE INFOCOM 2001*, pp. 1360–1369, 2001.

[11] P. Gupta, P.R. Kumar, "Critical Power for Asymptotic Connectivity in Wireless Networks", *Stochastic Analysis, Control, Optimization and Applications*, Birkhauser, Boston, pp. 547–566, 1998.

[12] P. Gupta, P.R. Kumar, "The Capacity of Wireless Networks", *IEEE Trans. Information Theory*, Vol. 46, n. 2, pp. 388–404, 2000.

[13] D.B. Johnson, D.A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks", *Mobile Computing*, Kluwer Academic Publishers, pp. 153–181, 1996.

[14] L.M. Kirousis, E. Kranakis, D. Krizanc, A. Pelc, "Power Consumption in Packet Radio Networks", *Theoretical Computer Science*, Vol. 243, pp. 289–305, 2000.

[15] V.F. Kolchin, B.A. Sevast'yanov, V.P. Chistyakov, *Random Allocations*, V.H. Winston and Sons, Washington D.C., 1978.

[16] J. Li, C. Blake, D.S.J. De Couto, H. Imm, R. Morris, "Capacity of Ad Hoc Wireless Networks", *Proc. ACM Mobicom 01*, pp. 61–69, 2001.

[17] Z. Lei, C. Rose, "Probability Criterion Based Location Tracking Approach for Mobility Management of Personal Communication Systems", *Proc. IEEE GLOBECOM 97*, pp. 977–981, 1997.

[18] S. Meguerdichian, F. Koushanfar, M. Potkonjak, M. Srivastava, "Coverage Problems in Wireless Ad Hoc Sensor Networks", *Proc. IEEE Infocom 2001*, pp. 1380–1387, 2001.

[19] S. Meguerdichian, S. Slijepcevic, V. Karayan, M. Potkonjak, "Localized Algorithms in Wireless Ad Hoc Networks: Location Discovery and Sensor Exposure", *Proc. ACM MobiHoc01*, pp. 106–116, 2001.

[20] A. Nasipuri, R. Castaneda, S.Das, "Performance of Multipath Routing for On-Demand Protocols in Mobile Ad Hoc Networks", *Mobile Networks and Applications*, Vol. 6, pp. 339–349, 2001.

[21] P. Panchapakesan, D. Manjunath, "On the Transmission Range in Dense Ad Hoc Radio Networks", *Proc. IEEE SPCOM 2001*, 2001.

[22] M.D. Penrose, "The Longest Edge of the Random Minimal Spanning Tree", *The Annals of Applied Probability*, Vol. 7, n. 2, pp. 340–361, 1997.

[23] M.D. Penrose, J.E. Yukich, "Central Limit Theorems for Some Graphs in Computational Geometry", *to appear in Annals of Applied Probability.*

[24] T.K. Philips, S.S. Panwar, A.N. Tantawi, "Connectivity Properties of a Packet Radio Network Model", *IEEE Trans. Information Theory*, Vol. 35, n. 5, pp. 1044–1047, 1989.

[25] P. Santi, D.M. Blough, F. Vainstein, "A Probabilistic Analysis for the Range Assignment Problem in Ad Hoc Networks", *Proc. ACM MobiHoc01*, pp. 212–220, 2001.

[26] J.M. Steele, "Growth Rates of Euclidean Minimal Spanning Trees with Power Weighted Edges", *Annals of Probability*, Vol. 16, pp. 1767–1787, 1988.

[27] Y. Xu, J. Heidemann, D. Estrin, "Geography-Informed Energy Conservation for Ad Hoc Routing", *Proc. ACM Mobicom 01*, pp. 70–84, 2001.

[28] M.M. Zonoozi, P. Dassanayake, "User Mobility Modeling and Characterization of Mobility Patterns", *IEEE Journal of Selected Areas in Comm.*, Vol. 15, n. 7, pp. 1239–1252, 1997.