An Evaluation of Connectivity in Mobile Wireless Ad Hoc Networks[✝]

Paolo Santi
Istituto di Informatica e Telematica
Area della Ricerca del CNR
Via G.Moruzzi, 1 – 56124 Pisa – Italy
santi@iit.cnr.it

Douglas M. Blough
School of Elec. and Comp. Eng.
Georgia Institute of Technology
Atlanta, GA 30332-0250 – USA
doug.blough@ece.gatech.edu

Abstract

We consider the following problem for wireless ad hoc networks: assume n nodes, each capable of communicating with nodes within a radius of r, are distributed in a d-dimensional region of side l; how large must the transmitting range r be to ensure that the resulting network is connected? We also consider the mobile version of the problem, in which nodes are allowed to move during a time interval and the value of r ensuring connectedness for a given fraction of the interval must be determined. For the stationary case, we give tight bounds on the relative magnitude of r, n and l yielding a connected graph with high probability in 1-dimensional networks, thus solving an open problem. The mobile version of the problem when d=2 is investigated through extensive simulations, which give insight on how mobility affect connectivity and reveal a useful trade-off between communication capability and energy consumption.

1 Introduction

Wireless ad hoc networks are networks where multiple nodes, each possessing a wireless transceiver, form a network amongst themselves via peer-to-peer communication. An ad hoc network can be used to exchange information between the nodes and to allow nodes to communicate with remote sites that they otherwise would not have the capability to reach. Wireless ad hoc networks are sometimes referred to as wireless multi-hop networks because, as opposed to wireless LAN environments, messages typically require multiple hops before reaching a gateway into the wired network infrastructure.

Sensor networks [8] are a particular class of wireless ad hoc networks in which there are many nodes, each containing application-specific sensors, a wireless transceiver, and a simple processor. Potential applications of sensor networks abound, e.g. monitoring of ocean temperature to enable more accurate weather prediction, detection of forest fires occurring in remote areas, and rapid propagation of traffic information from vehicle to vehicle, just to name a few. While the results in this paper apply to wireless ad hoc networks in general, certain aspects of the formulation are specifically targeted to sensor networks. For example, we assume nodes are randomly placed, which could result when sensors are distributed over a region from a moving vehicle such as an airplane. We are also concerned, in part, with minimizing energy consumption, which, although being an important issue in wireless ad hoc networks in general, is vital in sensor networks. Sensor nodes are typically battery-powered and, because replacing or recharging batteries is often very difficult or impossible, reducing energy consumption is the only way to extend network lifetime.

In many applications of wireless ad hoc networks, the nodes are mobile. This complicates analysis of network characteristics because the network topology is constantly changing in this situation. In this work, we consider networks both with and without mobility. We present analytical results that apply to networks without mobility and confine ourselves to simulation results for networks with mobility due to the intractability of analysis with existing mathematical methods.

Due to the relatively recent emergence of sensor networks, many fundamental questions remain unanswered. We address one of those questions, namely what are the conditions that must hold to ensure that a deployed network is connected initially and remains connected as nodes migrate? We address this question, and a number of related ones, in probabilistic terms, i.e. we evaluate the probabilities of various events related to network connectedness. The specific conditions we evaluate are how many nodes are required and what transmitting ranges must they have in order to establish a wireless ad hoc network with a particular property, e.g. connectedness. Determining an appropriate transmitting range for a given number of nodes is essential to minimize energy consumption since transmitting power is proportional to the square (or, depending on environmental

[✝] This research was supported in part by the National Science Foundation under Grant CCR-9803741.

0-7695-1597-5/02 $17.00 © 2002 **IEEE**

conditions, to a higher power) of the transmitting range. Our evaluation of required transmitting range is also useful in directing various 'topology control' protocols, which try to dynamically adjust transmitting ranges in order to minimize energy consumption at run time [6,9,10]. The question of how many nodes are necessary for a given transmitting range is very important for planning and design of wireless ad hoc networks when devices employ a fixed transceiver technology.

Our primary analytical result in this paper shows that a 1-dimensional network with nodes placed over a region of length l is connected if and only if the product of the number of nodes and the transmitting range is on the order of at least $l \log l$. This closes a gap between lower and upper bounds on this product that were established in an earlier paper [11]. Note that the 1-dimensional version of the problem does have important practical applications. The most notable such application is to cars on a freeway, which approximates a 1-dimensional region. An oft-cited potential use of mobile ad hoc networks is to have transmitters placed in cars that can transmit information about congestion or accidents to cars further back. By repeated relaying of such information, drivers far from the problem site can learn of the congestion and select an alternate route without waiting for a central notification system to learn of the event and post warning notices.

We also evaluate 2-dimensional networks with mobility through extensive simulations. We compare two different mobility models, the random waypoint model, which models intentional movement, and the drunkard model, in which movement is random. In both mobility models, we have included a parameter that 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.

The goal of our simulations is to study the relationship between the value of the transmitting range ensuring connected graphs in the stationary case and the values of the transmitting range ensuring connected graphs during some fraction of the operational time. In this paper, we focus on the transmitting ranges needed to ensure connectedness during 100%, 90% and 10% of the simulation time. These values are chosen as indicative of three different dependability scenarios that the ad hoc network must satisfy. We also consider the value of the transmitting range ensuring that the average size of the largest connected component is a given fraction of the total number of nodes in the network. The rationale for this investigation is that the network designer could be interested in maintaining only a certain fraction of the nodes connected, if this would result in significant energy savings. Further, considering that in many scenarios (e.g. wireless sensor networks) the cost of a node is very low, it could also be the case that dispersing twice as many nodes as needed and setting the transmitting ranges in such a way that half of the nodes remain connected is a feasible and cost-effective solution.

The results of our simulations have shown the somewhat surprising fact that, from a strictly statistical view of connectedness and connected component size, there are no major differences between the two mobility models. We also demonstrate that quite large reductions in transmitting range can be achieved if brief periods of disconnection are allowed and/or the network is allowed to operate with only a significant fraction of the nodes being connected. These results illustrate an energy vs. quality of communication trade-off that can be achieved in ad hoc networks, whereby the extent of communication capability can be somewhat reduced without great impact on the application and with the benefit of significantly reduced energy consumption. A final interesting result of our simulations shows that if about ½ or fewer of the nodes are mobile, then the network appears equivalent, in terms of statistical connectedness, to one without mobility.

The properties we study in this paper are akin to a simple form of availability for wireless ad hoc networks. Assuming that a network is "up" if all nodes are connected and "down" otherwise, then the percentage of time it is connected is an estimate of network availability. Since, in some applications, the network might be functional if at least a given fraction of nodes are connected, we also study the size of the largest connected component when the network is disconnected. For these applications, the percentage of time for which a sufficiently large number of nodes are connected is an availability estimate.

2 Preliminaries

A d-dimensional mobile wireless ad hoc network is represented by a pair $M_d=(N,P)$, where N is the set of nodes, with $|N|=n$, and P: $N\times T\rightarrow[0,l]^d$, for some $l>0$, is the *placement function*. The placement function assigns to every element of N and to any time $t\in T$ a set of coordinates in the d-dimensional cube of side l, representing the node's physical position at time t. The choice of limiting the admissible physical placement of nodes to a bounded region of $\mathbf{R}^d$ of the form $[0,l]^d$, for some $l>0$, is realistic and will ease the probabilistic analysis of Section 3. If the physical node placement does not vary with time, the network is said to be *stationary*, and function P can be redefined simply as P: $N\rightarrow[0,l]^d$.

In this paper, we assume that all the nodes in the network have the same transmitting range r. With this assumption, the *communication graph* of M_d induced at time t, denoted $G_M(t)$, is defined as $G_M(t)=(N,E(t))$, where the edge $(u,v)\in E(t)$ if and only if v is at distance at most r from u at time t. If $(u,v)\in E(t)$, node v is said to be a *neighbor* of u at time t. $G_M(t)$ corresponds to a *point graph* as defined in [12].

In the next section, we consider probabilistic solutions to the following problem for stationary ad hoc networks:

MINIMUM TRANSMITTING RANGE (MTR):
Suppose n nodes are placed in $[0,l]^d$; what is the minimum value of r such that the resulting communication graph is connected?

Given the number of nodes, minimizing r while maintaining a connected network is of primary importance if energy consumption is to be reduced. In fact, the energy consumed by a node for communication is directly dependent on its transmitting range. Further, a small value of r reduces the interferences between node transmissions, thus increasing the network capacity [5]. Observe that we could just as easily have stated the problem as one of finding the minimum number of nodes to ensure connectedness given a fixed transmitting range. This formulation is of primary importance in many dimensioning problems arising in the design of wireless ad hoc networks. For example, solving this problem would answer the following fundamental question to the system designer: for a given transmitter technology, how many nodes must be distributed over a given region to ensure connectedness with high probability? In fact, our solutions typically specify requirements on the product of n and r^d that ensures connectedness. These solutions can, therefore, be used to solve either MTR, as specified above, or the alternate formulation where the number of nodes is the primary concern.

It should be observed that the solution to MTR depends on the information we have about the physical node placement. If the node placement is known in advance, the minimum value of r ensuring connectedness can be easily determined. Unfortunately, in many realistic scenarios of ad hoc networks the node placement cannot be known in advance, for example because nodes are spread from a moving vehicle (airplane, ship or spacecraft). If nodes' positions are not known, the minimum value of r ensuring connectedness in all possible cases is $r \approx l\sqrt{d}$, which accounts for the fact that nodes could be concentrated at opposite corners of the placement region. However, this scenario appears to be very unlikely in most realistic situations. For this reason MTR has been studied in [1,11] under the assumption that nodes are distributed independently and uniformly at random in the placement region.

Observe that connectivity problems with formulations similar to MTR have also been studied in [4,7]. However, in these papers the deployment area is a fixed region (the unit disk in [4], or $[0,1]^2$ in [7]), and the number of nodes is increased to infinity. Thus, the asymptotic investigation is for networks with increasing node density, and is expected to be accurate in dense networks. On the contrary, the problem formulation used in this paper does not force the node density to asymptotically increase to infinity.

In the next section, we will improve the results of [1,11] for the case d=1 by means of a more accurate analysis of the conditions leading to disconnected communication graphs. The analysis will use some results of the occupancy theory [3], which are presented next.

The occupancy problem can be described as follows: assume we have C cells, and n balls to be thrown independently in the cells. The allocation of balls into cells can be characterized by means of random variables describing some property of the cells. The occupancy theory is aimed at determining the probability distribution of such variables as n and C grow to infinity (i.e., the *limit distribution*). The most studied random variable is the number of empty cells after all the balls have been thrown, which will be denoted $\mu(n,C)$ in the following. Under the assumption that the probability for any particular ball to fall into the i-th cell is $1/C$ for $i=1,\dots,C$ (*uniform allocation*), the following results have been proved[1]:

- $P(\mu(n,C)=0)=\sum_{i=0}^{C}\binom{C}{i}(-1)^i\left(1-\frac{i}{C}\right)^n$
- $E[\mu(n,C)]=C\left(1-\frac{1}{C}\right)^n$
- $Var[\mu(n,C)]=C(C-1)\left(1-\frac{2}{C}\right)^n+C\left(1-\frac{1}{C}\right)^n-C^2\left(1-\frac{1}{C}\right)^{2n}$,

where $E[\mu(n,C)]$ and $Var[\mu(n,C)]$ denote the expected value and the variance of $\mu(n,C)$, respectively. The asymptotic behaviors of $P(\mu(n,C)=k)$, $E[\mu(n,C)]$ and $Var[\mu(n,C)]$ depend on the relative magnitudes of n and C as they grow to infinity. The following theorem has been proved:

Theorem 1. *For every n and C, $E[\mu(n,C)]\le Ce^{-\alpha}$, where $\alpha=n/C$. Furthermore, if $n,C\to\infty$ in such a way that $\alpha=o(C)$, then:*

- $E[\mu(n,C)]=Ce^{-\alpha}-\frac{\alpha}{2}e^{-\alpha}+O\left(\frac{\alpha(\alpha+1)e^{-\alpha}}{C}\right)$
- $Var[\mu(n,C)]=Ce^{-\alpha}\left(1-(1+\alpha)e^{-\alpha}\right)+O\left(\alpha(1+\alpha)e^{-\alpha}\left(e^{-\alpha}+\frac{1}{C}\right)\right)$

Using the asymptotic formulas of Theorem 1, we can distinguish five different domains such that $n,C\to\infty$, for which the asymptotic distribution of the random variable $\mu(n,C)$ is different. These domains are:

- the *central domain* (CD for short), when $n=\Theta(C)$;

[1] All the results presented in this section are taken from [3].

- the *right-hand domain* (RHD for short), when $n=\Theta(C\log C)$;
- the *left-hand domain* (LHD for short), when $n=\Theta(\sqrt{C})$;
- the *right-hand intermediate domain* (RHID for short), when $n=\Omega(C)$ but $C\log C >> n$;[2]
- the *left-hand intermediate domain* (LHID for short), when $n=O(C)$ but $n>>\sqrt{C}$.

The following theorem describes the limit distribution of $\mu(n,C)$ in the different domains.

Theorem 2. *The limit distribution of the random variable $\mu(n,C)$ is:*
- *the normal distribution of parameters ($E[\mu(n,C)]$, $\sqrt{Var[\mu(n,C)]}$) in the CD, RHID and LHID;*
- *the Poisson distribution of parameter λ in the RHD, where $\lambda=\lim_{n,C\to\infty} E[\mu(n,C)]$.*

Furthermore, in the LHD the limit distribution of the random variable $\eta(n,C)=\mu(n,C)-(C-n)$ is the Poisson distribution of parameter ρ, where $\rho=\lim_{n,C\to\infty} Var[\mu(n,C)]$.

3 Probabilistic analysis of MTR for stationary networks

Consider the probability space $(\Omega_l,\mathcal{F}_l,P_l)$, where $\Omega_l=[0,l]$, $\mathcal{F}_l$ is the family of all closed subsets of Ω_l and P_l is a probability distribution on Ω_l. In this paper, we assume that P_l is the uniform distribution on Ω_l. Under this setting, nodes in N can be modeled as independent random variables uniformly distributed in $[0,l]$, which will be denoted $Z_1,\ldots,Z_n$.

We say that an event V_k, describing a property of a random structure depending on a parameter k, holds *asymptotically almost surely* (a.a.s. for short), if $P(V_k)\to 1$ as $k\to\infty$. In the following we consider the asymptotic behavior of the event CONNECTED$_l$ on the random structures $(\Omega_l,\mathcal{F}_l,P_l)$ as $l\to\infty$. Informally speaking, event CONNECTED$_l$ corresponds to all the values of the random variables $Z_1,..,Z_n$ for which the communication graph is connected.

The following upper bound on the magnitude of rn ensuring a.a.s. connectedness has been derived in [1].

Theorem 3. *Suppose n nodes are placed in [0,l] according to the uniform distribution. If $rn\in\Theta(l\log l)$ and $r>>1$, then the communication graph is a.a.s. connected.*

[2] Notation $f(x)<<g(x)$ (resp., $f(x)>>g(x)$) is used to denote the fact that $f(x)/g(x)\to 0$ (resp., ∞) as $x\to\infty$.

Observe that the constraint $r>>1$ in the statement of the theorem is not restrictive, since we are interested in investigating the magnitudes of r such that $1<<r<<l$.

In [11], a lower bound on the magnitude of rn ensuring a high probability of connectedness is derived by analyzing the probability of existence of an isolated node. In fact, the existence of an isolated node implies that the resulting communication graph (which is a point graph) is disconnected. However, the class of disconnected point graphs is much larger than the class of point graphs containing at least one isolated node. For this reason, the bounds established in [11] are not tight, and the gap between the lower and upper bounds on the magnitude of rn is in the order of $\log l$. In [1], it is conjectured that the upper bound stated in Theorem 3 is actually tight. This intuition has been experimentally confirmed by the results of extensive simulations [1,11]. In what follows, we prove that the conjecture stated in [1] is true for 1-dimensional networks. The result derives from a more accurate approximation of the class of disconnected point graphs, which is based on occupancy theory. This allows us to "close the gap", proving the tightness of the bound stated in Theorem 3.

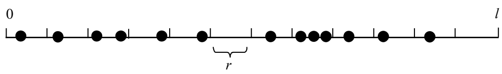

Figure 1. Node placement generating a disconnected communication graph.

In order to derive the lower bound, we consider the following subdivision of the placement region into cells. We assume that a line of length l is subdivided into $C=l/r$ segments of equal length r. With this subdivision, if there exists an empty cell c_i separating two cells c_{i-1},c_{i+1} that each contains at least one node, then the nodes in c_{i-1} are unable to communicate to those in c_{i+1}, and the resulting communication graph is disconnected (see Figure 1). The following lemma, whose immediate proof is omitted, establishes a sufficient condition for the communication graph to be disconnected.

Lemma 1. *Assume that n nodes are placed in [0,l], and that the line is divided into $C=l/r$ segments of equal length r. Assign to every cell c_i, for i=0,..,C-1, a bit b_i, denoting the presence of at least one node in the cell. Without loss of generality, assume $b_i=0$ if c_i is empty, and $b_i=1$ otherwise. Let $B=\{b_0...b_{C-1}\}$ be the string obtained by concatenating the bits b_i, for i=0,..,C-1. If B contains a substring of the form {10*1}, where 0* denotes that one or more 0s may occur, then the resulting communication graph is disconnected.*

Observe that the condition stated in Lemma 1 is sufficient but not necessary to produce a disconnected graph. In fact, there exist node placements such that B does not contain any substring of the form {10*1}, but the resulting communication graph is disconnected.

Let us denote with CONNECTED$_l$, DISCONNECTED$_l$, and E_l^{10*1} the events corresponding to all the values of the random variables Z_1,..Z_n such that the resulting communication graph is connected, disconnected, or a substring of the form {10*1} occurs in B, respectively. The subscript l indicates that we are considering these events in the case that the length of the line is l. Since CONNECTED$_l$=Ω_l-DISCONNECTED$_l$ and $E_l^{10*1} \subset$DISCONNECTED$_l$, it is immediate that a necessary condition for a.a.s. connectedness is that $\lim_{l\to\infty} P\left(E_l^{10*1}\right) = 0$.

In order to evaluate $\lim_{l\to\infty} P\left(E_l^{10*1}\right)$, we decompose the event E_l^{10*1} by conditioning on the disjoint events $\{\mu(n,C)=k\}$, for k=0,..,C; i.e.,

$$P\left(E_l^{10*1}\right) = \sum_{k=0}^{C} P\left(E_l^{10*1} \mid \{\mu(n,C)=k\}\right) \cdot P(\mu(n,C)=k) \qquad (1).$$

Observe that when l goes to infinity $P\left(E_l^{10*1}\right)$ is defined as the sum of an infinite number of non-negative terms $t_1, t_2, \ldots$. Clearly, if there exists at least one term $t_{\bar{k}}$ such that $\lim_{l\to\infty} t_{\bar{k}} = \varepsilon > 0$, then $\lim_{l\to\infty} P\left(E_l^{10*1}\right) \geq \varepsilon > 0$. In what follows, we prove that if l<<rn<<l log l and $\bar{k} = \lceil E[\mu(n,C)] \rceil$, then $\lim_{l\to\infty} t_{\bar{k}} = \varepsilon > 0$, thus implying that the resulting communication graph is not a.a.s. connected.

We start with a lemma that characterizes the asymptotic behavior of $P\left(E_l^{10*1} \mid \{\mu(n,C)=k\}\right)$ as l goes to infinity.

Lemma 2. *If 0<k<<C, then* $\lim_{l\to\infty} P\left(E_l^{10*1} \mid \{\mu(n,C)=k\}\right) = 1$.

Proof. *See Appendix.*

We now state the main theorem of this section.

Theorem 4. *Assume that l<<rn<<l log l. Then* $\lim_{l\to\infty} P\left(E_l^{10*1}\right) \geq \varepsilon > 0$.

Proof. *See Appendix.*

Combining the result stated in Theorem 4 with the bound of Theorem 3, we conclude this section with the following theorem.

Theorem 5. *Suppose n nodes are placed in [0,l] according to the uniform distribution, and assume 1<<r<<l. The communication graph is a.a.s. connected if and only if* $rn \in \Omega(l \log l)$.

The result stated in Theorem 5, for random distribution of nodes, should be compared to the transmitting ranges necessary with worst-case and best-case placements. To illustrate this, consider the case where the number of nodes is linear with the length of the line, l. In the worst-case, nodes are clustered at either end of the line and the transmitting range must be $\Omega(l)$ for the network to be connected. In the best-case placement, nodes are equally spaced at intervals of l/n, which in this case is a constant. Hence, a constant transmitting range is sufficient in the best case. Theorem 5's result yields a transmitting range of $\Omega(\log l)$ with random placement. Thus, there is a substantial reduction in transmitting range from the worst-case but also a significant increase compared to the best-case.

4 Evaluation of MTR for mobile networks

In this section, we consider the mobile version of MTR, which can be formulated as follows:

MINIMUM TRANSMITTING RANGE MOBILE (MTRM):
Suppose n nodes are placed in $[0,l]^d$*, and assume that nodes are allowed to move during a time interval [0,T]. What is the minimum value of r such that the resulting communication graph is connected during some fraction, f, of the interval?*

A formal analysis of MTRM is much more complicated than that of MTR and is beyond the scope of this paper. In this section, we study MTRM by means of extensive simulations. The goal is to study the relationship between the value of r ensuring connected graphs in the stationary case (denoted $r_{stationary}$) and the values of the transmitting range ensuring connected graphs during some fraction of the operational time. In this paper, we focus on the transmitting ranges needed to ensure connectedness during 100%, 90% and 10% of the simulation time (denoted r_{100}, r_{90} and r_{10}, respectively). These values are chosen as indicative of three different dependability scenarios that the ad hoc network must satisfy. In the first case, the network is used for safety-critical or life-critical applications (e.g., systems to detect physical intrusions in a home or business), and network connectedness during the entire operational time is a vital requirement. In this scenario, the potentially high price (expressed in terms of increased energy consumption) to be paid to keep the network always connected is a secondary issue. In the second case, temporary network disconnections can be tolerated, especially if this is counterbalanced by a significant decrease of the energy consumption with respect to the case of continuous connectedness. This scenario is plausible in many applications of wireless ad hoc networks, e.g. when the network is used to connect a squad of workers in an oil platform. In the latter case, the network stays disconnected most of the time, but temporary connection periods can be used to exchange data among nodes. This could be the case of wireless sensor networks [8] used for environmental monitoring [13], where environmental data (e.g., temperature, pressure, air pollution levels) are gathered by sensors, which periodically exchange these data with the other nodes in order to build a global view of the monitored region. In this setting, reducing energy consumption is the primary concern, and temporary connectedness is sufficient to ensure that the data sent by a sensor is eventually received by the other nodes in the network.

4.1 Simulation models

To generate the results of this section, we extended the simulator used in [1,11] for the stationary case by implementing two mobility models. The simulator distributes n nodes in $[0,l]^d$ according to the uniform distribution, then generates the communication graph assuming that all nodes have the same transmitting range r. Parameters r, n, l and d are given as input to the simulator, along with the number of iterations to run and the number, *#steps*, of mobility steps for each iteration. Setting *#steps*=1 corresponds to the stationary case. The simulator returns the percentage of connected graphs generated, the average size of the largest connected component (averaged over the runs that yield a disconnected graph) and the minimum size of the largest connected component. All of these parameters are reported with reference both to a single iteration (in this case, the averages are over all the mobility steps) and to all the iterations. In all simulations reported herein, we set d=2, as the two-dimensional setting is an appropriate model for many applications of wireless ad hoc networks.

Two mobility models have been implemented. The first model is the classical random waypoint model [2], and is used to model intentional movement: every node chooses uniformly at random a destination in $[0,l]^d$, and moves toward it 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. In the simulator, t_{pause} is expressed as the number of mobility steps for which the node must remain stationary. We have also included a further parameter in the model, namely the probability $p_{stationary}$ that a node remains stationary during the entire simulation time. Hence, only $(1-p_{stationary})n$ nodes (on the average) will move. Introducing $p_{stationary}$ 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.

The second mobility model resembles a drunkard-like (i.e., non-intentional) motion. Mobility is modeled using parameters $p_{stationary}$, p_{pause} and m. Parameter $p_{stationary}$ is defined as above. Parameter p_{pause} is the probability that a node remains stationary at a given step. This parameter accounts for heterogeneous mobility patterns, in which nodes may move at different times. Intuitively, the higher is the value of p_{pause}, the more heterogeneous is the mobility pattern. However, values of p_{pause} close to 1 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 disk of radius m 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.

4.2 Simulation results for increasing system size

The first set of simulations was aimed at investigating the value of the ratio of r_{100} (respectively, of r_{90} and r_{10}) to $r_{stationary}$ for values of l ranging from 256 to 16384. We also considered the largest value r_0 of the transmitting range that yields no connected graphs. In both mobility models, n was set to $\sqrt{l}$. The value of $r_{stationary}$ is obtained from the simulation results for the stationary case reported in [1,11], while those for r_{100}, r_{90}, r_{10} and r_0 are averaged over 50 simulations of 10000 steps of mobility each.

First, we considered the random waypoint model, with parameters set as follows: $p_{stationary}$=0, v_{min}=0.1, v_{max}=0.01l, and t_{pause}=2000. This setting models a moderate mobility scenario, in which all the nodes are moving, but their velocity is rather low. The effect of different choices of the mobility parameters on the values of r_{100}, r_{90} and r_{10} is studied in the next sub-section. The values of the ratios are reported in Figure 2. Figure 3 reports the same graphic obtained for the drunkard model, with $p_{stationary}$=0.1, p_{pause}=0.3 and m=0.01l. This is also a moderate mobility scenario, but more heterogeneous than the other: a small percentage of the nodes remain stationary, and mobile nodes are stationary for 30% of the simulation time (on average).

The graphics show the same qualitative behavior: as l increases, the ratio of the different transmitting ranges for mobility to $r_{stationary}$ tends to increase, and this increasing behavior is more pronounced for the case of r_{100}. However, even when the system is large, a modest increase to $r_{stationary}$ (about 21% in the random waypoint and about 25% in the drunkard model) is sufficient to ensure connectedness during the entire simulation time. Comparing the results for the two mobility models, we can see somewhat higher values of the ratios for the drunkard model, especially for the case of r_{100}. This seems to indicate that more homogeneous mobility patterns help in maintaining connectedness. However, it is surprising that the results for the two mobility models are so similar. This indicates that it is more the existence of mobility rather than the precise details of how nodes move that is significant, at least as far as network connectedness is concerned.

It should also be observed that r_{90} is far smaller than r_{100} (about 35-40% smaller) in both mobility models,

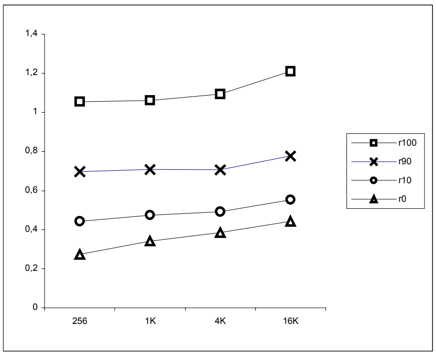

Figure 2. Values of $r_x/r_{stationary}$ for increasing values of l in the random waypoint model.

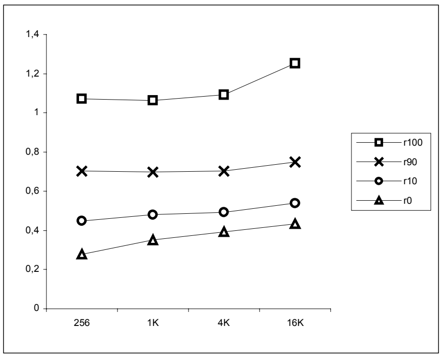

Figure 3. Values of $r_x/r_{stationary}$ for increasing values of l in the drunkard model.

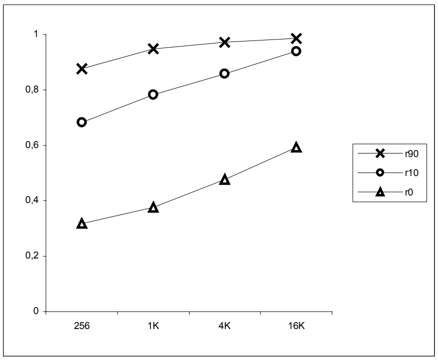

Figure 4. Average size of the largest connected component (expressed as a fraction of n) for increasing values of l in the random waypoint model.

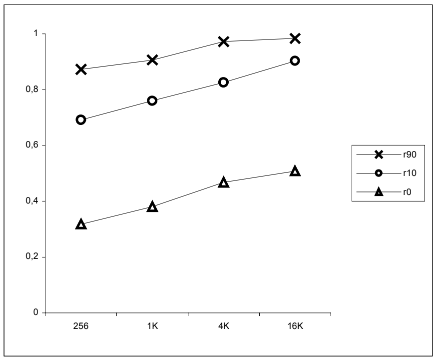

Figure 5. Average size of the largest connected component (expressed as a fraction of n) for increasing values of l in the drunkard model.

independently of the system size. Hence, substantial energy savings can be achieved under both models if temporary disconnections can be tolerated. When the requirement for connectedness is only 10% of the operational time, the decrease in the transmitting range is about 55-60%, enabling further energy savings. However, if r is reduced to about 25% to 40% of $r_{stationary}$, the network becomes disconnected during the entire simulation time.

The average size of the largest connected component when the transmitting range is set to r_{90}, r_{10} and r_0 was also investigated. Simulation results are displayed in Figures 4 and 5. Once again, the graphics show very similar behaviors: the ratio of the average size of the largest connected component to n increases as l increases. When the transmitting range is set to r_{90} and l is sufficiently large, this ratio is very close to 1 (about 0.98 in both mobility models). This means that during the short time in which the network is disconnected, a vast majority of its nodes forms a large connected component. Hence, on the average disconnection is caused by only a few isolated nodes. This fact is confirmed by the plots for r_{10}: even when the network is disconnected most of the time, a large connected component (of average size about $0.9n$ for large values of l) still exists. However, if the transmitting range is further decreased to r_0, the size of the largest connected component drops to about $0.5n$.

We also considered the value of the transmitting range ensuring that the average size of the largest connected component is $0.9n$, $0.75n$ and $0.5n$, respectively. The corresponding values of the transmitting range are denoted r_{l90}, r_{l75} and r_{l50}. The mobility parameters and n were set as above. The rationale for this investigation is that the network designer could be interested in maintaining only a certain fraction of the nodes connected, if this would result in significant energy savings. Further, considering that in many scenarios (e.g. wireless sensor networks) the cost of a node is very low, it could also be the case that dispersing twice as many nodes as needed and setting the transmitting ranges in such a way that half of the nodes remain connected is a feasible and cost-effective solution.

The value of the ratio of r_{l90}, r_{l75} and r_{l50} to $r_{stationary}$ for increasing values of l in the random waypoint model is shown in Figure 6. Simulation results have shown that while $r_{l90}/r_{stationary}$ tends to decrease with increasing values of l, converging to about 0.52, the ratios $r_{l75}/r_{stationary}$ and

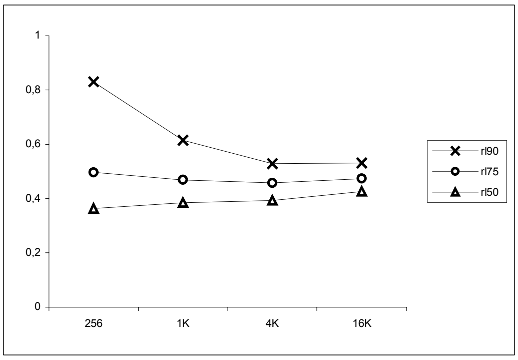

Figure 6. Values of the ratio r_{l90}, r_{l75} and r_{l50} to $r_{stationary}$ for l ranging from 256 to 16384 in the random waypoint model.

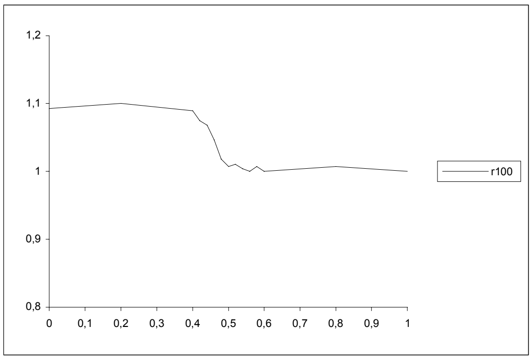

Figure 7. Value of $r_{100}/r_{stationary}$ for different values of $p_{stationary}$ in the random waypoint model.

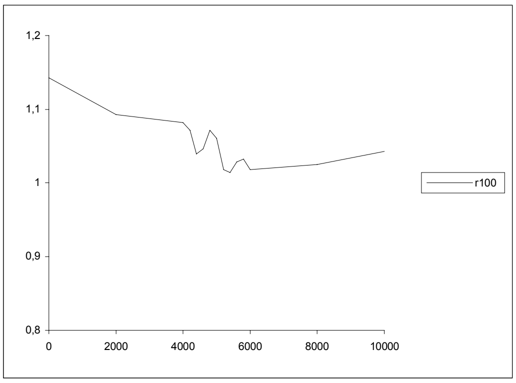

Figure 8. Value of $r_{100}/r_{stationary}$ for values of t_{pause} ranging from 0 to 10000 in random waypoint model.

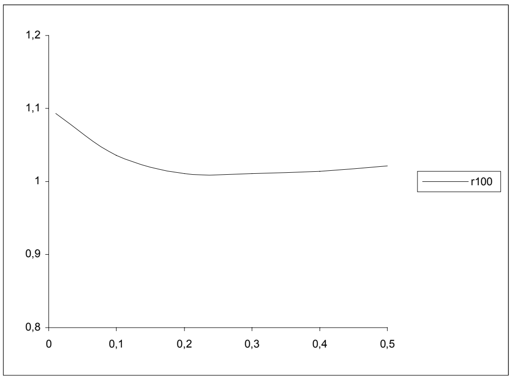

Figure 9. Value of $r_{100}/r_{stationary}$ for values of v_{max} ranging from $0.01l$ to $0.5l$ in random waypoint model.

$r_{l50}/r_{stationary}$ are almost independent of l. In particular, $r_{l75}/r_{stationary}$ is about 0.46 and $r_{l50}/r_{stationary}$ is about 0.4. Further, the relative differences between the three ratios decrease for increasing value of l. This indicates that, while for small networks (few nodes distributed in a relatively small region) the energy needed to maintain 90% of the nodes connected is significantly higher than that required to connect 50% of the nodes (r_{l50} is less than half of r_{l90} for l=256), for large networks the savings are not as great if the requirement for connectivity is only 50% of the nodes (r_{l50} is 20% smaller than r_{l90} for l=16384).

4.3 Simulation results for different mobility patterns

A second set of simulations was done to investigate the effect of different choices of the mobility parameters on the value of r_{100}. We considered the random waypoint model with l=4096 and $n=\sqrt{l}=64$. The default values of the mobility parameters were set as above, i.e. $p_{stationary}$=0, v_{min}=0.1, v_{max}=0.01l, and t_{pause}=2000. Then, we varied the value of one parameter, leaving the others unchanged.

Figure 7 reports the value of r_{100} for values of $p_{stationary}$ ranging from 0 (no stationary nodes) to 1 (corresponding to the stationary case) in steps of 0.2. Simulation results show a sharp drop of r_{100} in the interval 0.4-0.6: for $p_{stationary}$=0.4, r_{100} is about 10% larger than $r_{stationary}$, while for $p_{stationary}$=0.6 and for larger values of $p_{stationary}$ we have $r_{100} \approx r_{stationary}$. To investigate this drop more closely, we performed further simulations by exploring the interval 0.4-0.6 in steps of 0.02. As shown in Figure 7, there is a distinct threshold phenomenon: when the number of stationary nodes is about $n/2$ or higher, the network can be regarded as practically stationary from a connectedness point of view. This result is very interesting, since it seems to indicate that a certain number (albeit a rather large fraction) of stationary nodes would significantly increase network connectedness. With more than $n/2$ mobile nodes, the network quickly becomes equivalent to one in which all nodes are mobile.

The effect of t_{pause} and of the velocity on r_{100} is shown in Figures 8 and 9. Increasing values of t_{pause} tend to decrease the value of r_{100}, although the trend is not as pronounced as in the case of $p_{stationary}$. A threshold phenomenon seems to exist in the interval 4000-6000 in this case also. However, further simulations in this interval have shown that, although the trend can be observed, no sharp threshold actually exists. We believe that the rationale for this is the following: while the value of $p_{stationary}$ has a direct impact on the "quantity of mobility" (which can be informally understood as the percentage of stationary nodes with respect to the total number of nodes), the effect of the pause time is not so direct. In fact, in the random waypoint model the "quantity of mobility" depends heavily on the node destinations, which are chosen uniformly at random: even if the pause time is long and the velocity is moderate, a node could be "mobile" for a long time if its destination is very far from its initial location. So, an increased pause time tends to render the system more stationary, but in a much less direct way than $p_{stationary}$.

As shown in Figure 9, the value of r_{100} is almost independent of the value of v_{max}: except for low velocities (v_{max} below $0.1l$), r_{100} is slightly above $r_{stationary}$. This surprising result could be due to the apparently counterintuitive fact that the “quantity of mobility” is only marginally influenced by the value of v_{max}, and a larger value of v_{max} tends to decrease the “quantity of mobility”. In fact, the larger is v_{max}, the more likely it is that nodes arrive quickly at destination and remain stationary for t_{pause}=2000 steps.

5. Conclusions

In this paper, we considered a connectivity problem in the case of both stationary and mobile wireless ad hoc networks. For the stationary case, we have derived tight bounds on the magnitude of r, n and l ensuring connectedness with high probability for 1-dimensional networks. Our bounds improve on existing results, and prove a conjecture stated in a previous paper. We have also investigated the mobile version of the problem for 2-dimensional networks through extensive simulation. We implemented two motion patterns to model both intentional and non-intentional movements, and we simulated 2-dimensional networks of different sizes and using different mobility parameters. Simulation results have shown that consistent energy savings can be achieved if temporary disconnections can be tolerated or if connectedness must be ensured only for a large fraction of the nodes. Regarding the influence of mobility patterns, simulation results have shown that connectedness is only marginally influenced by whether motion is intentional or not, but it is rather related to the “quantity of mobility”, which can be informally defined as the percentage of stationary nodes with respect to the total number of nodes. For example, when about $n/2$ nodes are static, the network can be regarded as stationary from a connectivity point of view. Further investigation in this direction is needed, and is a matter of ongoing research.

References

[1] D.M. Blough, P. Santi, "The Random Point Graph Model for Ad Hoc Networks and its Application to the Range Assignment Problem", *Tech. Rep. IMC-B4-01-05*, Istituto di Matematica Computazionale del CNR, Pisa - Italy, Dec. 2001.

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

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

[4] P. Gupta, and P. R. Kumar, “Critical Power for Asymptotic Connectivity in Wireless Networks”, *Stochastic Analysis, Control, Optimization and Applications*, (W.M. McEneany, G. Yin, and Q. Zhang, eds.), Birkhauser, Boston, 1998, pp. 547 – 566.

[5] P. Gupta, P.R. Kumar, “The Capacity of Wireless Networks”, *IEEE Trans. Inf. Theory*, vol. 46, n. 2, pp. 388-404, March 2000.

[6] L. Li, J. H. Halpern, P. Bahl, Y. Wang, R. Wattenhofer, “Analysis of a Cone-Based Distributed Topology Control Algorithm for Wireless Multi-hop Networks”, *Proc. ACM Symp. on Principles of Distributed Computing (PODC)*, pp. 264-273, August 2001.

[7] T.K. Philips, S.S. Panwar, A.N. Tantawi, “Connectivity Properties of a Packet Radio Network Model”, *IEEE Trans. Inf. Theory*, vol. 35, n. 5, pp. 1044-1047, Sept. 1989.

[8] G. J. Pottie, W. J. Kaiser, “Wireless Integrated Network Sensors”, *Communications of the ACM*, vol. 43, no. 5, pp. 51- 58, May 2000.

[9] R. Ramanathan, R. Rosales-Hain, “Topology Control of Multihop Wireless Networks using Transmit Power Adjustment”, *Proc. IEEE Infocom 2000*, pp. 404 – 413, 2000.

[10] V. Rodoplu, T. H. Meng, “Minimum Energy Mobile Wireless Networks”, *IEEE J. Selected Areas in Comm.*, vol. 17, n. 8, pp. 1333-1344, Aug.1999.

[11] P. Santi, D. M. Blough, and F. Vainstein, “A Probabilistic Analysis for the Range Assignment Problem in Ad Hoc Networks”, *Proc. ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Long Beach, CA, October 2001, pp. 212-220.

[12] A. Sen, and M. L. Huson, “A New Model for Scheduling Packet Radio Networks”, *Proc. IEEE Infocom ’96*, pp. 1116 – 1124, 1996.

[13] D. C. Steere, A. Baptista, D. McNamee, C. Pu, J. Walpole, “Research Challenges in Environmental Observation and Forecasting Systems”, *Proc. ACM MOBICOM 2000*, pp. 292-299, 2000.

Appendix

Proof of Lemma 2.

Consider the complementary event of E_l^{10*1}, i.e. $E_l^1 = \Omega_l - E_l^{10*1}$. It can be easily seen that E_l^1 corresponds to all the values of the random variables $Z_1,..Z_n$ such that the 1-bits in B are consecutives. Given the hypothesis of independence of the random variables $Z_1,..Z_n$, when exactly k cells out of C are empty (i.e., k bits in B are 0), $P\left(E_l^1 \mid \{\mu(n,C)=k\}\right)$ corresponds to the ratio of all configurations of (n-k) consecutive 1-bits over all possible configurations of k 0-bits in C positions, i.e.:

$$P\left(E_l^1 \mid \{\mu(n,C)=k\}\right)=\frac{k+1}{\binom{C}{k}}.$$

Since $C=l/r$ and $r<<l$, we have:

$$\lim_{l\to\infty} P\left(E_l^{10^*1} \mid \{\mu(n,C)=k\}\right)=1-\lim_{l\to\infty} P\left(E_l^1 \mid \{\mu(n,C)=k\}\right)=1-\lim_{C\to\infty}\frac{k+1}{\binom{C}{k}}$$

We can rewrite the last limit as:

$$\lim_{C\to\infty}\frac{k+1}{\binom{C}{k}}=\lim_{C\to\infty}\frac{(k+1)!}{C(C-1)\ldots(C-k+1)}$$

Since $k<<C$, we have:

$$\lim_{C\to\infty}\frac{(k+1)!}{C(C-1)\ldots(C-k+1)}=\lim_{C\to\infty}\frac{(k+1)!}{C^k}$$

Taking the logarithm, we obtain:

$$\lim_{C\to\infty}\ln\frac{(k+1)!}{C^k}=\lim_{C\to\infty}\ln(k+1)!-k\ln C$$
$$=\lim_{C\to\infty}k\ln k-k\ln C=\lim_{C\to\infty}k(\ln k-\ln C)$$

Since $0<k<<C$, we conclude that $\lim_{C\to\infty}k(\ln k-\ln C)=-\infty$, hence:

$$\lim_{C\to\infty}\frac{k+1}{\binom{C}{k}}=0,$$

and the Lemma is proved.

❑

Proof of Theorem 4.

By Equation (1) and Lemma 2, it is sufficient to show that there exists a value $\bar{k}$ such that:

- $0<\bar{k}<<C$, and
- $\lim_{l\to\infty}P\left(\mu(n,C)=\bar{k}\right)\geq\varepsilon>0$.

Consider $\bar{k}=\lceil E[\mu(n,C)]\rceil$. From Theorem 1, we have that

$$\bar{k}\approx Ce^{-\alpha}-\frac{\alpha}{2}e^{-\alpha}+O\left(\frac{\alpha(\alpha+1)e^{-\alpha}}{C}\right), \qquad (2)$$

where $\alpha=n/C=rn/l$.

Since $l<<rn<<l\log l$, we have $1<<\alpha=f(l)<<\log l$, and (2) can be rewritten as

$$\bar{k}\approx\frac{l}{re^{f(l)}}, \text{ with } 0<\frac{1}{r}<<\bar{k}<<\frac{l}{r}=C,$$

hence the first condition is satisfied.

Observe that the condition $l<<rn<<l\log l$ implies that $C<<n<<C\log C$, i.e. we are in the RHID. By Theorem 2, it follows that the limit distribution of $\mu(n,C)$ as n,C go to infinity is the normal distribution of parameters ($E[\mu(n,C)]$, $\sqrt{Var[\mu(n,C)]}$). By Theorem 1, $Var[\mu(n,C)]$ can be rewritten as

$$Var[\mu(n,C)]=Ce^{-\alpha}\left(1-(1+\alpha)e^{-\alpha}\right)+O\left(\alpha(1+\alpha)e^{-\alpha}\left(e^{-\alpha}+\frac{1}{C}\right)\right)\approx\frac{l}{re^{f(l)}}\left(1-\frac{1+f(l)}{e^{f(l)}}\right)$$

Hence, we have:

$$P\left(\mu(n,C)=\bar{k}\right)\approx\frac{1}{\sqrt{2\pi Var[\mu(n,C)]}}\approx\sqrt{\frac{re^{f(l)}}{2\pi\cdot l}}$$

Let us choose $r=\delta\frac{l}{e^{f(l)}}$, for some $0<\delta\leq 2\pi$. Observe that this choice of r is consistent with the hypothesis $1<<r<<l$, since we have $1<<\frac{l}{e^{f(l)}}<<l$. With this choice of r, we can write the limit as follows:

$$\lim_{l\to\infty}P\left(\mu(n,C)=\bar{k}\right)=\lim_{l\to\infty}\sqrt{\frac{re^{f(l)}}{2\pi\cdot l}}=\sqrt{\frac{\delta}{2\pi}}=\varepsilon>0,$$

and the theorem is proved.

❑