Investigating Upper Bounds on Network Lifetime Extension for Cell-Based Energy Conservation Techniques in Stationary Ad Hoc Networks*

Douglas M. Blough
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta GA 30332-0250 - USA
doug.blough@ece.gatech.edu

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

ABSTRACT

Cooperative cell-based strategies have been recently proposed as a technique for extending the lifetime of wireless ad hoc networks, while only slightly impacting network performance. The effectiveness of this approach depends heavily on the node density: the higher it is, the more consistent energy savings can potentially be achieved. However, no general analyses of network lifetime have been done either for a base network (one without any energy conservation technique) or for one using cooperative energy conservation strategies. In this paper, we investigate the lifetime/density tradeoff under the hypothesis that nodes are distributed uniformly at random in a given region, and that the traffic is evenly distributed across the network. We also analyze the case where the node density is just sufficient to ensure that the network is connected with high probability. This analysis, which is supported by the results of extensive simulations, shows that even in this low density scenario, cell-based strategies can significantly extend network lifetime.

Categories and Subject Descriptors

C.2.1 [**Computer Systems Organization**]: Computer-Communication Networks—*Network Architecture and Design: Wireless Communication*

General Terms

Theory, performance

Keywords

Wireless ad hoc networks, network lifetime, energy conservation, cooperative strategies, occupancy theory

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

MOBICOM'02 September 23–26, 2002, Atlanta, Georgia, USA.

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 usually multi-hop networks because, as opposed to wireless LAN environments, messages typically require multiple hops before reaching a gateway into the wired network infrastructure.

Since nodes in a wireless ad hoc network are battery-powered, energy conservation is important to extend the functional lifetime of both individual nodes and the network. Much of the prior work on energy conservation in wireless ad hoc networks has focused on energy-conserving routing protocols, e.g. [4, 5, 12, 21]. However, the maximum energy savings, and hence lifetime extension, that can be achieved through routing algorithm optimization could be quite limited. This is because the energy used by standard wireless interfaces just to sense the channel can be nearly as much as that used in receiving a message and about 60% of the energy used in transmitting [8].

A few prior works have proposed more aggressive strategies for energy conservation in which some nodes' transceivers are completely shut down for some period of time [6, 14, 22]. This type of approach has the potential to reduce energy consumption dramatically and, therefore, to increase network lifetime significantly. However, no general analyses of network lifetime have been done either for a base network (one without any energy conservation techniques) or for one using an aggressive energy conservation technique.

In this paper, we analyze quantitatively the increase in network lifetime that results from using an idealized version of the cell-based energy conservation technique described in [22]. Indeed, what is investigated here is the potentiality of the class of energy conserving techniques based on a subdivision of the deployment region into non-overlapping cells, under the assumption that nodes are distributed uniformly at random. To this purpose, we will make a series of assumptions, aimed at simplifying the analysis on one hand, and at studying the achievable lifetime extension under the best possible conditions on the other hand. Our analysis can thus be used as an assessment framework for cell-based energy conservation techniques, similarly to the MERIT framework

for routing protocol assessment presented in [7]. An example of application of our assessment framework to the GAF protocol presented in [22] will be described in Section 6.

As expected, our analysis has shown that the lifetime of a network employing a cell-based energy conservation technique depends on the density with which nodes are distributed throughout the region where the network is deployed. In particular, the analysis has shown that when the node density is large enough to ensure that the resulting network is connected with high probability, the network lifetime is very likely to be extended by a multiplicative factor of at least h, with respect to the case where no energy conservation is used. The multiplicative factor h is primarily a function of the region's dimension. The results of our simulations have shown that h is about 3.5 for one-dimensional networks and 1.8 for two-dimensional networks. Furthermore, our analysis suggests that cell-based cooperative strategies have the potential to scale well with node density: the scaling is more than linear in the one-dimensional case and almost linear in the two-dimensional case.

Our analytical results are based on the application of occupancy theory to the distribution of nodes over a region. These results are supported by extensive simulations, which validate our analysis. Both the analytical results and the simulations demonstrate quite clearly the potential benefits of aggressive cell-based energy conservation techniques applied to wireless ad hoc networks.

2. DEFINITION OF NETWORK LIFETIME

2.1 Discussion

Nodes in a wireless ad hoc network are usually powered by a limited capacity battery. As nodes' batteries are drained and they stop functioning, the wireless ad hoc network will eventually cease to be usable. Informally speaking, we refer to the length of the time that the network operates prior to becoming unusable as the *network lifetime*.

A formal definition of network lifetime is not straightforward and may depend on the application scenario in which the network is used. In the literature, network lifetime has often been defined as the time for the first node to die (e.g., in [4, 5, 12, 21]), or as the time for a certain percentage of network nodes to die (as in [22]). Alternatively, network lifetime has been defined in terms of the packet delivery rate [6] or in terms of the number of alive flows [3], thus accounting for the "quality of communication" the alive nodes achieve.

We believe that the definitions above do not satisfactorily capture the intuition behind the concept of network lifetime, i.e. its "being usable". On one hand, referring only to the number of nodes fails to capture the operational aspect of the network: if a few nodes in strategic positions die, the network could become disconnected, thus impairing dramatically its functionality. Conversely, defining network lifetime in terms of the time for the first node to die is often pessimistic, since it is very likely that the surviving nodes remain connected, thus not impairing the network functionality. However, a definition of network lifetime expressed solely in terms of the "quality of communication" is not sufficient either. For example, if a large number of sensors in a given region of a wireless sensor network suddenly die because of a catastrophic event, the surviving sensors are very likely to remain connected, but the network functionality (which, in this specific case, should also be expressed in terms of the amount of the region covered by the network) is impaired.

The discussion above outlines that a good definition of network lifetime should refer to the capability of the network to provide the services it was designed for, and hence depends on the application scenario. An example of a definition in this sense can be found in [1], where the lifetime of a sensor network is defined as the time to the first loss of coverage[1]. However, this definition does not take connectivity into account, and a disconnected sensor network is most likely unusable.

Summarizing, in a very broad sense we can distinguish the following scenarios:

- *ad hoc networks*, whose primary goal is to provide connectivity to the nodes participating in the network;
- *sensor networks*, in which the requirement for connectivity must be complemented with the requirement for coverage; i.e., all the monitored region (or at least a significant percentage of it) must be covered by the sensors.

In the first case, the (minimal) service that the network must provide is connectedness, hence network lifetime should be defined as the time to the first disconnection. This definition can be refined with the further requirement that a minimum number of nodes remain alive. This accounts for the fact that a network composed by, say, 20% of the initial nodes, although connected, is very likely to be unable to provide the services for which it was designed.

In the second scenario, the definition of network lifetime should account also for coverage, and can thus be expressed as the time to the first loss of coverage or connectedness. Also in this case, we can impose the further requirement on the number of alive nodes.

Both definitions can be weakened, declaring the network dead when the size of the largest connected component or the percentage of the region covered drop below a certain threshold. We call these properties fractional connectivity and fractional coverage, respectively. This accounts for different dependability requirements the network must satisfy: if it is used for safety-critical applications, connectedness and coverage are vital requirements, and even a limited loss of coverage (or connectedness) cannot be tolerated. However, in less critical applications a small percentage of disconnected units (or of the region uncovered) can be tolerated, especially if this is counterbalanced by an extended lifetime.

Note that in the discussion above we have implicitly assumed that the nodes are stationary. If nodes are mobile, a network which is disconnected at time t_1 may become connected at time $t_2 > t_1$. In the case of sensor networks, the network coverage may display a similar behavior. Hence, defining lifetime in mobile networks is much more complicated than in the stationary case and is beyond the scope of this paper.

[1]The definition of coverage usually is based on the sensing range of nodes. In a simplified but widely used model, all the nodes sense a circular area of radius r_s, and the monitored region R is *covered* if every point of R is a distance of at most r_s from at least one sensor.

2.2 Formal definition

The network is composed by a set V of nodes, with $|V| = n$. Each node is equipped with a radio transceiver and is powered by a limited capacity battery. We assume that all the nodes in the network are homogeneous, i.e. equipped with devices with the same features. Hence, all the nodes have the same transmitting range and batteries with the same initial capacity.

Nodes are placed in a d-dimensional region R called the *deployment region*, with $d = 1, 2, 3$. For the sake of simplicity, we assume that R is of the form $[0, l]^d$, where l is the *length* of a side of the region. Given the node positions in R and the transmitting range r, the *communication graph* is defined as $G = (V, E)$, where edge (u, v) is in E if and only if $d(u, v) \leq r$, and $d(u, v)$ denotes the Euclidean distance between u and v. Observe that, since all the nodes have the same transmitting range, the communication graph is undirected.

We are now ready to give a general definition of network lifetime for the cases of general ad hoc networks and sensor networks.

Definition 1 (Ad hoc networks lifetime). *Let $G(t) = (V(t), E(t))$ be the communication graph of the ad hoc network at time t, where $V(t)$ is the set of alive nodes at time t. Assume that $G(0)$ is connected, and denote with $n(t)$ the cardinality of $V(t)$, with $n = n(0)$. The network lifetime is defined as the minimum between t_1 and t_2, where t_1 is the time it takes for the cardinality of the largest connected component of $G(t)$ to drop below $c_1 \cdot n(t)$, t_2 is the time it takes for $n(t)$ to drop below $c_2 \cdot n$, and $0 \leq c_1, c_2 \leq 1$.*

Definition 2 (Sensor networks lifetime). *Let $G(t) = (V(t), E(t))$ be the communication graph of the sensor network at time t, where $V(t)$ is the set of alive nodes at time t. Assume that $G(0)$ is connected and covers the deployment region $R = [0, l]^d$, and denote with $n(t)$ the cardinality of $V(t)$, with $n = n(0)$. The network lifetime is defined as the minimum between t_1, t_2 and t_3, where t_1 is the time it takes for the cardinality of the largest connected component of $G(t)$ to drop below $c_1 \cdot n(t)$, t_2 is the time it takes for $n(t)$ to drop below $c_2 \cdot n$, t_3 is the time it takes for the volume covered to drop below $c_3 \cdot l^d$, and $0 \leq c_1, c_2, c_3 \leq 1$.*

The definitions above are very general, and can be reduced to most existing definitions by appropriately choosing the values for c_1, c_2 and c_3. For example, in the case of general ad hoc networks setting $c_1 = 0$ and $c_2 = 1$ corresponds to defining lifetime as the time it takes for the first node to die, while setting $c_1 = 1$ and $c_2 = 0$ corresponds to defining lifetime as the time to network disconnection.

In the following, we will restrict our attention to the case of general ad hoc networks, leaving the analysis of sensor networks lifetime for future research.

3. AN IDEALIZED MODEL FOR CELL-BASED TECHNIQUES

Consider a wireless ad hoc network composed by n nodes, where all the nodes are equipped with batteries with the same initial capacity and have the same transmitting range r. The energy consumption for an individual node is dependent on how much traffic is routed through it and so it depends on the routing algorithm and traffic model. Several energy conserving routing algorithms have been proposed in the literature [4, 5, 12, 21]. However, the maximum energy savings, and hence lifetime extension, that can be achieved through routing algorithm optimization could be quite limited. This is because the energy used by standard wireless interfaces just to sense the channel can be nearly as much as that used in receiving a message and about 60% of the energy used in transmitting [8]. Thus, to achieve order of magnitude reductions in energy consumption, more aggressive measures, which we call *cooperative strategies*, should be considered.

Cooperative strategies are based on the following idea: assume that a given set S of nodes provides a functionality F to the rest of the system; instead of keeping all the nodes in S operative, a representative node u can be selected, and the remaining nodes can be turned off in order to save energy. The representative node selection is obtained as the result of a negotiation protocol executed by nodes in S, which is repeated when u dies, or after a certain wake-up time. This way, considerable energy savings can be potentially achieved. Observe that the exact definition of the cooperative strategy depends on the kind of functionality we are interested in. For example, F could be defined as the capability of nodes in S to relay messages on behalf of the remaining nodes, without compromising network connectivity (i.e., nodes are equivalent from the point of view of a routing protocol). In the case of sensor networks, F can be alternately defined as the capability of sensing a given sub-region of the deployment region R.

Examples of cooperative strategies can be found in [6, 22]. In [6], the authors present a coordination algorithm, called SPAN, aimed at reducing power consumption while preserving both the network capacity and connectivity. SPAN adaptively elects coordinators from all nodes in the network, which are left active, while non-coordinator nodes are shut down. The coordination algorithm is transparent to the routing protocol, and can be integrated in the IEEE 802.11 MAC layer. In [22], the deployment region R is divided into an appropriate number of non-overlapping cells, with the property that all the nodes in the same cell are equivalent from the routing protocol point of view. Nodes in the same cell elect a representative, which is left active, while the transceivers of non-representative nodes are shut down. Periodically, the representative election phase is repeated to balance power consumption and to deal with mobility.

A similar aggressive approach to reducing power consumption has been presented in [14], where the authors define a transport level protocol for shutting down and restarting the network interface. However, this work is intended for a mobile host communicating with a single base station and, therefore, does not address the network-wide issues that must be dealt with in a wireless ad hoc network.

The effectiveness of a cooperative strategy depends heavily on the node density. Intuitively, if node density is low, almost all the nodes must stay up all the time, and no energy saving can be achieved. Considering the overhead required for coordination of nodes, the actual network lifetime could actually be reduced with respect to the case where no cooperative strategy is used. Conversely, if node density is high, consistent energy savings (and, consequently, extension of network lifetime) can be achieved. This behavior is displayed by the GAF protocol of [22], while the energy savings achieved by SPAN does not increase with node density.

This is due to the fact that the overhead required for coordination with SPAN tends to "explode" with node density, and thus counterbalances the potential savings achieved by the increased density. For this reason, in the following we will focus our attention mainly on the GAF strategy.

We remark that our investigation can be directly applied to any cooperative strategy based on the subdivision of the deployment region into non-overlapping cells. To date, GAF is the only representative of this class of protocols, but, given the effectiveness of this approach, new cell-based energy conserving protocols are likely to be implemented. Thus, our analysis can be used to compare the energy savings achieved by a cell-based cooperative strategy with the "best possible" savings achievable in optimal conditions. Although in a less direct way, our analysis can also be used to assess general cooperative strategies that rely on node density to increase lifetime. For example, the scalability analysis of Section 5.2 shows that for two-dimensional networks, network lifetime has the potential to increase almost linearly with node density. Although this analysis is cell-based, it is reasonable to assume that any good cooperative strategy should display a similar behavior.

Observe that the positive effect of an increased node density on network lifetime could be counterbalanced by its detrimental effect on the network capacity. In fact, it is known that, at least in the stationary case, the network capacity does not scale with node density, and the end-to-end throughput achievable at each node goes to 0 as the density increases [11]. Furthermore, increasing node density entails a higher network cost. On the other hand, node density cannot be too low, since otherwise network connectivity would be impaired. Hence, the tradeoff between node density, network lifetime, and capacity/cost must be carefully evaluated. As a first step in this direction, in the following we investigate the relation between the expected benefit of the utilization of cooperative strategies and the node density, under the following simplifying hypotheses:

a1. nodes are distributed uniformly and independently at random in $R = [0, l]^d$, with $d = 1, 2, 3$;

a2. nodes are stationary;

a3. F is defined as the capability of nodes in S to relay messages on behalf of the remaining nodes, without compromising network connectivity. To this end, R is divided into non-overlapping d-dimensional cells of equal side $\frac{r}{2\sqrt{d}}$ [2][2]. The total number of cells is then $N = \frac{k_d l^d}{r^d}$, where $k_d = 2^d d^{d/2}$;

a4. we consider an ideal cooperative strategy, in which the overhead needed to coordinate nodes amongst themselves is zero. Hence, the energy savings derived in the following can be seen as the best possible a cooperative strategy can achieve;

a5. node density is sufficient to achieve connectedness with high probability;

a6. network lifetime is defined in terms of connectedness, i.e. setting $c_1 = 1$ and $c_2 = 0$ in Definition 1;

a7. the traffic is balanced over all cells;

a8. the energy consumed by other components of a node is negligible compared to the energy consumption of its transceiver.

Observe that, by assumption *a7*, all the cells are subject to the same load. If this load must be handled by a single node, it will die at the *baseline time* T; however, if a cell contains h nodes, the load can be evenly divided among them, and the last node in the cell will die at time hT. Hence, a lower bound to the network lifetime can be obtained by evaluating the probability distribution of the minimum number of nodes in a cell, and occupancy theory [13] can be brought to bear on the problem.

4. A LOWER BOUND TO NETWORK LIFETIME FOR THE IDEALIZED APPROACH

In this section we will use the standard notation regarding the asymptotic behavior of functions, which we recall. Let f and g be functions of the same parameter x. We have:

- $f(x) = O(g(x))$ if there exist constants C and x_0 such that $f(x) \leq C \cdot g(x)$ for any $x \geq x_0$;
- $f(x) = \Omega(g(x))$ if $g(x) = O(f(x))$;
- $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$. In this case, we also use the notation $f(x) \approx g(x)$;
- $f(x) = o(g(x))$ if $\frac{f(x)}{g(x)} \to 0$ as $x \to \infty$;
- $f(x) \ll g(x)$ or $g(x) \gg f(x)$ if $f(x) = o(g(x))$.

The probability distribution of the minimum number of nodes in a cell can be evaluated using some results of the occupancy theory [13], which studies properties of the random independent allocations of n balls into N urns[3] when $n, N \to \infty$. Let $\eta(n, N)$ be the random variable denoting the minimum number of nodes in a cell. The form of the limit distribution (i.e., of the probability distribution of $\eta(n, N)$ when $n, N \to \infty$) depends on the asymptotic behavior of the ratio $\frac{\alpha}{\ln N}$, where $\alpha = \frac{n}{N}$. The following theorem holds [13]:

THEOREM 1. *If* $\frac{\alpha}{\ln N} \to 1$ *as* $n, N \to \infty$ *and* $h = h(\alpha, N)$ *is chosen so that* $h < \alpha$ *and* $N p_h(\alpha) \to \lambda$, *where* $p_h(\alpha) = \frac{\alpha^h}{h!} e^{-\alpha}$ *and* λ *is a positive constant, then:*

- $P(\eta(n, N) = h) \to 1 - e^{-h}$
- $P(\eta(n, N) = h + 1) \to e^{-h}$

Theorem 1 states that if $\frac{\alpha}{\ln N} \to 1$ as $n, N \to \infty$, then $\eta(n, N)$ is either h or $h + 1$ asymptotically almost surely[4] (a.a.s. for short), where h is such that $N p_h(\alpha) \to \lambda$, for some positive constant λ. Similar results presented in [13] determine the limit distribution of $\eta(n, N)$ when $\frac{\alpha}{\ln N} \to x$, for some

[2]This ensures that a node in one cell can communicate with all nodes in the complete neighborhood of cells surrounding it. Note that the side of the cell as defined here is slightly different from that used in the GAF protocol [22], which ensures that nodes residing in a cell can communicate with all the nodes in the upper, lower, left and right cell. However, this slight difference does not impair the validity of our analysis for the GAF protocol.

[3]For consistency, in the following we will use the words *node* and *cell* instead of *ball* and *urn*, respectively.

[4]We say that an event E_m, describing a property of a random structure depending on a parameter m, holds *asymptotically almost surely* if $P(E_m) \to 1$ as $m \to \infty$.

$x > 1$, or when $\frac{\alpha}{\ln N} \to \infty$. However, for our purposes it is sufficient to note that, denoting by $\eta_1(n,N)$, $\eta_x(n,N)$ and $\eta_\infty(n,N)$ the value of $\eta(n,N)$ for the three asymptotic cases, it is $\eta_1(n,N) \leq \eta_x(n,N) \leq \eta_\infty(n,N)$ a.a.s. This follows immediately by the fact that in the three asymptotic cases a strictly increasing number of nodes are distributed into the same number of cells.

Observe that Theorem 1 gives very precise information on the asymptotic value of $\eta(n,N)$, but gives no explicit value of h. In the following, we derive the value of h under the hypothesis that the network is a.a.s. connected.

It is known [2, 19, 20] that a sufficient condition for a.a.s. connectedness when n nodes with transmitting range r are distributed uniformly at random in $R = [0,l]^d$ is that $r^d n = \Theta(l^d \log l)$, and that this condition is also necessary for $d = 1$. However, these results give us only the magnitude of the density which is sufficient to ensure a.a.s. connectedness. A more precise result is stated in the following Theorem, the proof of which is reported in the Appendix.

THEOREM 2. *Assume that n nodes, each with transmitting range r, are distributed uniformly and independently at random in $R = [0,l]^d$, for $d = 1,2,3$, and assume that $r^d n = k l^d \ln l$ for some constant $k > 0$, with $r = r(l) \ll l$ and $n = n(l) \gg 1$. If $k > d \cdot k_d$, or $k = d \cdot k_d$ and $r = r(l) \gg 1$, then $\lim_{l\to\infty} P_{conn}(l) = 1$, where $k_d = 2^d d^{d/2}$ and $P_{conn}(l)$ denotes the probability that the communication graph is connected.*

PROOF. See Appendix. □

Observe that the conditions on the magnitude of $r = r(l)$ and $n = n(l)$ in the statement of Theorem 2 are not restrictive. In fact, if $r = \Omega(l)$, then every node is able to transmit directly to most of the other nodes, and connectedness is ensured independently of n. The condition $n = n(l) \gg 1$ is a straightforward consequence of the first condition, since otherwise the probability of connectedness would be negligible.

Theorem 2 specifies the values of the multiplicative constant which are sufficient to ensure a.a.s. connectedness, and is a generalization of the results presented in [10, 16, 17, 18], where similar evaluations for the case of n nodes distributed in a region of constant side, or under the assumption that nodes are distributed in R with a given Poisson density δ, are performed.

We are now ready to evaluate the value of h in the statement of Theorem 1. Given the result of Theorem 2 and assumption *a3*, we have $r^d n \geq d \cdot k_d l^d \ln l$ and $N = \frac{k_d l^d}{r^d}$. In order to satisfy the hypothesis of the theorem, we must have $\lim_{n,N\to\infty} \frac{\alpha}{\ln N} = 1$, which, given the hypotheses $r = r(l) \ll l$ and $n = n(l) \gg 1$, is equivalent to $\lim_{l\to\infty} \frac{\alpha}{\ln N} = 1$. We have:

$$\lim_{l\to\infty} \frac{\alpha}{\ln N} = \lim_{l\to\infty} \frac{\frac{k}{k_d} \ln l}{\ln \frac{k_d l^d}{r^d}} = \lim_{l\to\infty} \frac{\ln l^{k/k_d}}{\ln \frac{k_d l^d}{r^d}} \qquad (1)$$

Setting $k = d \cdot k_d$ we obtain:

$$(1) = \lim_{l\to\infty} \frac{\ln l^d}{\ln \frac{k_d l^d}{r^d}} = 1$$

when $r^d = \Theta(1)$.

Let now consider the expression $N p_h(\alpha)$, which can be rewritten as

$$\frac{k_d \left(\ln l^d\right)^h}{r^d h!} \qquad (2)$$

Since $r^d = \Theta(1)$, (2) $\approx \frac{\left(\ln l^d\right)^h}{h!}$. Taking the logarithm, we have:

$$\ln \frac{\left(\ln l^d\right)^h}{h!} = h \ln \left(\ln l^d\right) - \ln h! \approx h \ln \left(\ln l^d\right) - h \ln h$$

It follows that $N p_h(\alpha) \to \lambda$, for some positive constant λ, if and only if $h \approx \ln l^d$.

Based on the discussion above, we can conclude this section with the following theorem:

THEOREM 3. *Assume that n nodes with transmitting range r are distributed uniformly and independently at random in $R = [0,l]^d$, for $d = 1,2,3$, and assume that $r^d n \geq d \cdot k_d l^d \ln l$, where $k_d = 2^d d^{d/2}$. If a cooperative strategy is used to alternately shut down "routing equivalent" nodes, then $P(NL_l \geq hT) \to 1$ as $l \to \infty$, where NL_l is the random variable denoting network lifetime and $h = d(1-\epsilon)\ln l$, for any constant ϵ such that $0 < \epsilon < 1$.*

PROOF. The proof follows by Theorem 1 and by the discussion above, and by observing that:

- when $k = dk_d$ and $r^d \gg 1$, $\lim_{n,N\to\infty} \frac{\alpha}{\ln N} = \infty$;
- when $k > dk_d$, $\lim_{n,N\to\infty} \frac{\alpha}{\ln N} = \infty$ if $r^d \gg 1$, and $\lim_{n,N\to\infty} \frac{\alpha}{\ln N} = x$, for some $x > 1$, otherwise.

□

Theorem 3 gives a lower bound to an ad hoc network's lifetime under the hypothesis that the network is a.a.s. connected. The value of h depends on the dimension of the network and on the side of the deployment region. For example, if $r = 250$ and $l = 5000$, network lifetime is extended at least 2.99 times when $d = 1$, 5.99 times when $d = 2$, and 8.98 times when $d = 3$.[5] However, these bounds are mainly of theoretical interest: in fact, simulations reported in [20] have shown that, while the sufficient condition for a.a.s. connectedness is tight for $d = 1$, it becomes looser for two and three-dimensional networks. For this reason, we have performed extensive simulations to evaluate the actual network lifetime when the minimal condition for connectedness holds. The results of these simulations are reported in the next section.

It should also be observed that in the optimal case, i.e. when the n nodes are evenly distributed into the N cells, the network lifetime is exactly $d \cdot k_d \ln l$. The result stated in Theorem 3 is very important, since it states that, in the case of nodes distributed uniformly at random, the network lifetime differs from the optimal at most by a constant factor, namely a factor of $k_d - 1 + \epsilon$, for some $0 < \epsilon < 1$. This factor is dependent on the dimension d of the network: it equals $1 + \epsilon$ when $d = 1$, but it is rather large ($\approx 40.5 + \epsilon$) when $d = 3$. The quality of this result is confirmed by the simulation results presented in the next section.

Finally, we observe that Theorem 3 can be used to lower bound the network lifetime also when c_2 in Definition 1 is

[5] In calculating h, l is measured as a multiple of r.

greater than 0. In fact, Theorem 3 actually lower bounds the time for the first node to die under the hypothesis that the network is a.a.s. connected.

5. SIMULATION RESULTS

In this section we present the results of extensive simulations. The goal of these simulations was to:

- evaluate the actual network lifetime with the minimal node density that achieves connectedness with high probability;
- investigate the density vs. lifetime tradeoff;
- investigate the effect of the parameter c_2 in the definition of lifetime;
- validate the theoretical analysis presented in the previous section.

The simulator distributes n nodes in $R = [0, l]^d$ according to the uniform distribution and generates the communication graph assuming that all the nodes have the same transmitting range r. The deployment region R is then divided into the appropriate number of cells, and the number of nodes in every cell is calculated. Next, the actual network lifetime is evaluated: all the nodes contained in the cell(s) with the minimum number of nodes are declared dead, and the communication graph is modified accordingly. If it remains connected and the number of alive nodes is at least $c_2 \cdot n$, the process is iterated until one of the two conditions is impaired. The current version of the simulator checks for network connectivity rather than fractional connectivity. Thus, all results in this section are for $c_1 = 1$.

The input parameters of the simulator are n, l, d, r and parameter c_2 of Definition 1, along with the number *iter* of iterations to run. The output parameters are the percentage of connected graphs generated, the minimum number of nodes in a cell after the initial node placement, and the number of nodes in the cell(s) whose death "kills" the network (the *terminating cell(s)*). This number is the most important output data of the simulation, since it determines the network lifetime extension with respect to the baseline time. The simulator also reports the number of nodes alive at network death and, in case parameter c_2 is not 0, whether the network died because of disconnection or because of too few nodes remaining alive. The results returned by the simulator are averaged over the *iter* simulation runs.

One important point to note is that the simulator accepts integer values for l and r, while the cell size in general is not an integer. This means that smaller cells on the boundary of the deployment region may exist. While values of l and r generating equally-sized cells can be chosen in one-dimensional networks, this is not possible for two and three-dimensional networks, due to the $\sqrt{d}$ term in the expression for the cell size. Hence, for these networks we chose combinations of the l and r values producing boundary cell sizes that are nearly equal to the inner cell size. Since the boundary cells are slightly smaller than those in the center, the results reported by the simulator could be slightly different from those that would be obtained if all cells were of equal size. However, this inaccuracy is not critical, both because it is quite small and because it can only cause our simulator to underestimate the minimum number of nodes in a cell and the network lifetime.

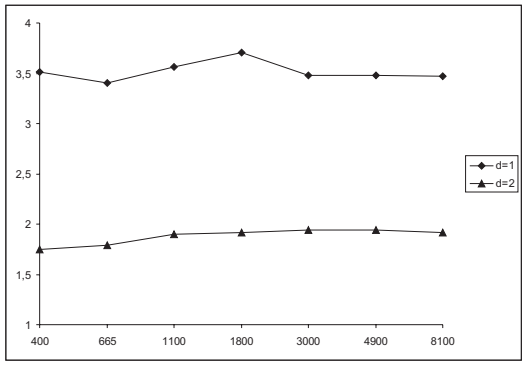

Figure 1: Average number of nodes in the terminating cell for regions of increasing side l.

We simulated one-dimensional and two-dimensional networks with sides of length l ranging from about 400 to about 8100. For one-dimensional networks, we chose combinations of l and r generating equally-sized cells, where $r \approx \frac{l}{\ln^2 l}$. For two-dimensional networks, we chose l and r in order to generate nearly equally-sized cells, where $r \approx \frac{l}{\ln l}$. For each combination of l and r, we experimentally calculated the minimum number n_{min} of nodes to be distributed in order to obtain 100% of connected graphs in 1000 simulation runs[6]. This number gave us the minimal node density $\delta_{min} = \frac{n_{min}}{l^d}$ which was used as the baseline for our simulations.

We performed separate sets of simulations to investigate each of the above mentioned issues. Simulation results, all averaged over 1000 runs, are reported in the following subsections.

5.1 Network lifetime with minimal density

The first set of simulations was aimed at evaluating the network lifetime with the minimal density δ_{min}. In this situation, the minimum number of nodes in a cell is very small: it is 1.3–1.5 for $d = 1$, and it is 0 for $d = 2$. However, this number only represents a lower bound to the actual network lifetime.

Simulation results, which are reported in Figure 1, show that even in this apparently unfavorable situation the network lifetime is increased significantly on average. The average number of nodes in the terminating cell(s) (i.e., the cell that "kills" the network) is about 3.5 for $d = 1$ and about 1.8 for $d = 2$, and is far above the minimum number of nodes in a cell. This means that network lifetime is extended on the average by 250% in one-dimensional and by 80% in two-dimensional networks[7]. The larger extension for $d = 1$ than for $d = 2$ is a consequence of the fact that in one-dimensional networks the minimal cell density is higher than in two-dimensional networks: when $n = n_{min}$, the average number of nodes in a cell is about 7 when $d=1$,

[6] n_{min} was found to range from 543 to 1180 in one-dimensional networks, and from 284 to 797 in the two-dimensional case.

[7] Since the number of nodes in a cell is an integer, saying that the network lifetime is extended on the average of 80% means the following: for some networks (those in which there is only one node in the terminating cell) there is no extension, while for other networks the lifetime is doubled, or tripled, and so on. The average of these extensions is 80%.

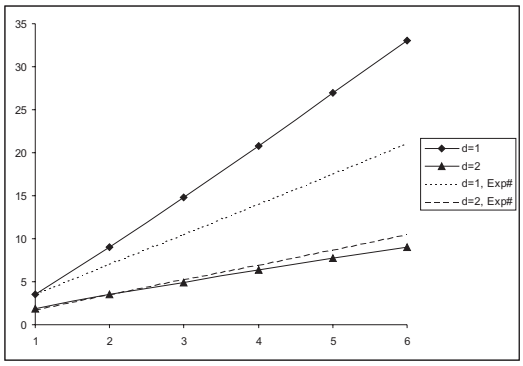

Figure 2: Average number of nodes in the terminating cell with increasing density for one and two-dimensional networks.

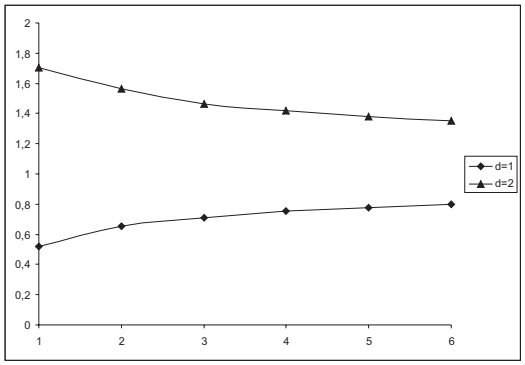

Figure 3: Ratio between the average number of nodes in the terminating cell and the expected number of nodes in a cell for increasing density.

while it is about 1 when $d = 2$. This happens because in one-dimensional networks, two consecutive empty cells separating non-empty cells cause disconnection, while in two-dimensional networks connectivity can be ensured by nodes that "go around the hole" in one dimension. Hence, when $d=1$ more nodes must be distributed in order to avoid the occurrence of empty cells, which can be tolerated when $d=2$.

5.2 Density vs. lifetime

Since for the minimum connecting density and $d=2$, there was a non-negligible percentage of networks that experienced *no* lifetime extension, we also investigated how much lifetime increases as the node density is increased beyond the minimum level. We distributed $\beta \cdot n_{min}$ nodes into one-dimensional and two-dimensional networks of sides $l = 400$ and $l = 1100$, with $\beta = 1, \ldots, 6$. The results of these simulations are shown in Figure 2.

The average number of nodes in the terminating cell is only marginally influenced by the length of a side of the deployment region. For this reason, Figure 2 reports only the data for the case $l = 1100$. The figure also reports the expected number of nodes in the terminating cell, which is simply β times the average number of nodes in the terminating cell when $n = n_{min}$. As shown in the figure, the actual number is far above the expected number of nodes in the terminating cell for $d=1$, while it is slightly below it for $d=2$. This means that, if the node density is $\beta\delta_{min}$ (with $\beta \geq 2$) and T_{min} is the network lifetime when $n = n_{min}$, the average network lifetime is far above βT_{min} in one-dimensional networks, while it is slightly below βT_{min} in two-dimensional networks.

Another important aspect to consider is the relation between the number of nodes in the terminating cell and the expected number of nodes in a cell, which accounts for the case in which the nodes are perfectly distributed among the cells. Figure 3 plots the ratio of these numbers for increasing density. We considered one and two-dimensional networks, with $l = 1100$. As it is seen, this ratio tends to converge towards the perfectly-distributed case in both one and two-dimensional networks. Quite surprisingly, this ratio is above 1 in two-dimensional networks, meaning that a uniform node distribution is better than when nodes are perfectly distributed. This is due to the fact that, as discussed above, two-dimensional networks are often connected even if many cells are empty. Hence, many cells must die before the network becomes disconnected.

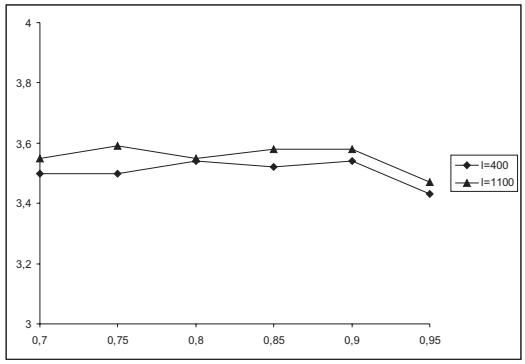

Figure 4: Average number of nodes in the terminating cell for different values of c_2 in one-dimensional networks.

5.3 Varying c_2

In previous simulations we had $c_1 = 1$ and $c_2 = 0$. In this set of simulations, we have investigated the effect of a different value of c_2 on the network lifetime. We recall that $c_2 > 0$ forces there to be a minimum fraction of the original nodes remaining alive in order for the entire network to be considered alive. We considered one-dimensional and two-dimensional networks of sides $l = 400$ and $l = 1100$.

Figure 4 refers to one-dimensional networks, and reports the average number of nodes in the terminating cell when $n = n_{min}$ nodes are distributed, with c_2 ranging from 0.7 to 0.95 in steps of 0.05. As it is seen, there is a significant drop in network lifetime only for $c_2 = 0.95$. This is due to the fact that, in these conditions, the average number of alive nodes when the network becomes disconnected is $0.926n$ when $l = 400$ and $0.931n$ when $l = 1100$. Thus, below $c_2 = 0.95$, the connectedness requirement dominates, while above that, the c_2 condition is significant.

Figure 5 refers to two-dimensional networks. In this case, we distributed $n = 2 \cdot n_{min}$ nodes, since the average number of alive nodes at network disconnection when $n = n_{min}$ is about $0.34n$. As it is seen, in this case the drop in network

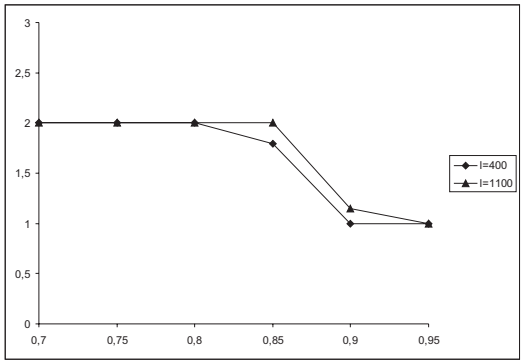

Figure 5: Average number of nodes in the terminating cell for different values of c_2 in two-dimensional networks.

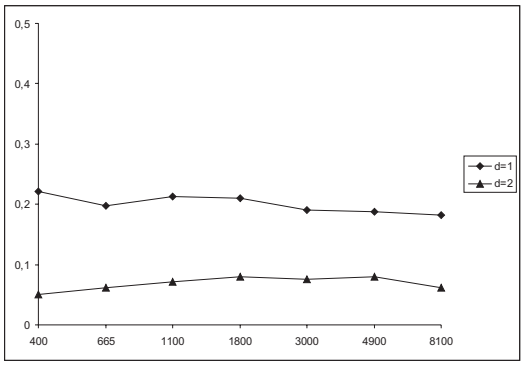

Figure 6: Average value of the ratio ρ for regions of increasing side l.

lifetime is for c_2 ranging from 0.8 to 0.9.

5.4 Validating the theoretical analysis

The goal of the last set of simulations was to validate the theoretical result of Theorem 3. In essence, Theorem 3 states that, if the condition for a.a.s. connectedness holds, the minimum and the expected number of nodes in a cell have the same magnitude $\Theta(\ln l)$. Hence, the ratio ρ between the minimum and the expected number of nodes in a cell should remain constant as the the length l of a side of the deployment region increases. Furthermore, since h in the statement of the Theorem differs from $d \cdot k_d \ln l$ by a factor that increases with the dimension of the network, the value of ρ should be larger for $d = 1$ than for $d = 2$.

As shown in Figure 6, this behavior is actually displayed by the results of our simulation. The plot for $d = 1$ refers to the case in which n_{min} nodes are distributed, and shows that the value of the constant is approximately 0.2. In two-dimensional networks, more nodes must be distributed, since the minimum number of nodes in a cell for $n = n_{min}$ was always 0. Observe that Theorem 3 still applies, since it holds not only for the minimal density ensuring a.a.s. connectedness, but also for larger densities. The plot reported in the figure is obtained by setting $n = 3 \cdot n_{min}$, and the value of the constant is about 0.06. This value is far below 0.2, which, on the other hand, was obtained with a lower node density.

6. GAF ASSESSMENT

As an example of application of our assessment framework to a realistic case, consider the GAF protocol of [22]. The authors have investigated the node density vs. lifetime tradeoff for a network using GAF, where network lifetime is defined as the time for the packet delivery rate to drop below 80%. Nodes, each with a nominal transmitting range of 250 meters, are distributed uniformly at random in a 1500×300 meters area, and traffic is generated by sources spreading the messages randomly among 10 traffic nodes, which are not part of the simulation (i.e., their energy consumption is not considered). The hypotheses of [22] are slightly different from ours: network lifetime as defined in [22] is clearly related to connectivity, but it is different from the time to network disconnection. Furthermore, the aspect ratio of the deployment region is in between the one and the two-dimensional case, although it is probably closer to the 1-d case (in fact, the transmitting range almost covers the entire smaller side of the region). Nevertheless, a fruitful comparison between the performance evaluation presented in [22] and our analysis can be done. When the node density is increased 4 times (with respect to a minimal density which, although not explicitly stated, is sufficient to ensure connectedness), the lifetime of a network using GAF is extended by approximately a factor of 4 (in the case that nodes are stationary). This can be compared with the results of our simulations, which have shown that the average lifetime extension is slightly below 4 times the baseline lifetime when $d = 2$, and it is about 20 times for $d = 1$. Another parameter to be considered is the cell density, i.e. the expected number of nodes in a cell. This parameter is 4.4 for the highest density scenario of [22]. With similar values of this parameter, the lifetime extension in two-dimensional networks in our ideal analysis is about 6.34 times. Hence, GAF shows very good scalability, but it leaves some room for further improvement.

7. CONCLUDING REMARKS

In this paper we have investigated the node density vs. network lifetime tradeoff for a cell-based energy conservation technique in wireless ad hoc networks.

We have presented a lower bound to network lifetime that holds under the assumption that nodes are distributed uniformly at random in a given region, that node density is just sufficient to guarantee connectedness with high probability, and that the traffic is evenly distributed among the cells. This theoretical result has been validated by means of extensive simulations, which have complemented the theoretical (qualitative) with a quantitative analysis. In particular, the results of our simulations have shown that even in the minimal density scenario network lifetime can be extended significantly with respect to the case where no energy conservation strategy is used. However, while the amount of this extension for one-dimensional networks is large, in the two-dimensional case the "best possible" extension is only about 80%. This indicates that in this scenario even more aggressive techniques could be considered.

For example, we could consider *probabilistic* instead of *deterministic* equivalence: in our analysis and in the GAF protocol, it is assumed that all the nodes in a cell are within the transmitting range of all the nodes in the neighboring cells. Hence, it is sufficient to leave one node up in every cell to

ensure connectivity. As an alternative, we could use slightly larger cells, where nodes are within the range of almost all the nodes in neighboring cells. This means that nodes in the same cells are not "routing equivalent", but "routing equivalent with a given probability", which depends on the length of a cell's side. The benefit of this technique would be an increased cell density given the same node density, which could result in an increased lifetime.

If weaker requirements on network connectivity (e.g., at least 90% of nodes in the largest connected component) can be tolerated, an alternative strategy to save further energy would be to reduce the node transmitting range. In fact, results presented in [2] for two-dimensional networks have shown that halving the minimal transmitting range for connectedness still yields a largest connected component containing about 90% of the network nodes. However, reducing the transmitting range would reduce the cell density (given the same node density), and a detrimental effect on the benefit of cooperative cell-based strategies could occur. Hence, the adoption of a similar approach should be carefully evaluated. Observe that one difficulty in comparing the two approaches stems from the fact that the analysis presented in this paper refers to a baseline time, but if we consider different transmitting ranges, the baseline time would be different also. Thus, we should express network lifetime referring to an absolute time axis, rather than to a baseline lifetime.

Regarding the "lifetime scalability", the results of our simulations have shown that, as node density increases, network lifetime increases more than linearly in one-dimensional networks, and almost linearly in two-dimensional networks. Hence, cell-based cooperative strategies have the potential to "scale well" with node density.

Finally, further investigation is needed to gain insights on how mobility affects the effectiveness of cell-based cooperative strategies. The intuition as well as the theoretical and experimental results presented in [9, 22] seem to indicate that mobility has a positive effect, since the average number of nodes in a particular cell over time tends to match the expected number of nodes in a cell overall. However, the results of our simulation have shown that in two-dimensional networks a uniform node distribution outperforms the "ideal case" of evenly distributed nodes. Whether the same behavior is displayed in presence of mobility or not is an interesting issue, and is a matter of ongoing research.

8. ACKNOWLEDGMENTS

The authors wish to thank Stefano Basagni for helpful discussions on the concept of network lifetime.

9. REFERENCES

[1] M. Bhardwaj, T. Garnett, and A. Chandrakasan. Upper bounds on the lifetime of sensor networks. In *Proc. IEEE International Conference on Communications*, pages 785–790, 2001.

[2] D. Blough and 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, Pisa, 2001.

[3] T. Brown, H. Gabow, and Q. Zhang. Maximum flow-life curve for a wireless ad hoc network. In *Proc. ACM MobiHoc 01*, pages 128–136, 2001.

[4] J. Chang and L. Tassiulas. Routing for maximum system lifetime in wireless ad hoc networks. In *Proc. 37th Annual Allerton Conference on Communication, Control and Computing*, 1999.

[5] J. Chang and L. Tassiulas. Energy conserving routing in wireless ad hoc networks. In *Proc. IEEE INFOCOM 2000*, pages 22–31, 2000.

[6] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. Span: An energy-efficient coordination algorithm for topology maintainance in ad hoc wireless networks. In *Proc. ACM Mobicom 01*, pages 85–96, 2001.

[7] A. Farago and V. Syrotiuk. Merit: A unified framework for routing protocol assessment in mobile ad hoc networks. In *Proc. ACM Mobicom 01*, pages 53–60, 2001.

[8] L. Feeney and M. Nilson. Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In *Proc. IEEE INFOCOM 2001*, pages 1548–1557, 2001.

[9] M. Grossglauser and D. Tse. Mobility increases the capacity of ad hoc wireless networks. In *Proc. IEEE INFOCOM 2001*, pages 1360–1369, 2001.

[10] P. Gupta and P. Kumar. Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications*, Birkhauser - Boston:547–566, 1998.

[11] P. Gupta and P. Kumar. The capacity of wireless networks. *IEEE Trans. Information Theory*, 46(2):388–404, 2000.

[12] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient routing protocols for wireless microsensor networks. In *Proc. 33rd Hawaii International Conference on System Sciences*, pages 1–10, 2000.

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

[14] R. Kravets and P. Krishnan. Power management techniques for mobile communication. In *Proc. ACM Mobicom 99*, pages 24–35, 1999.

[15] E. Palmer. *Graphical Evolution*. John Wiley and Sons, New York, 1985.

[16] P. Panchapakesan and D. Manjunath. On the transmission range in dense ad hoc radio networks. In *Proc. IEEE SPCOM 2001*, 2001.

[17] T. Philips, S. Panwar, and A. Tantawi. Connectivity properties of a packet radio network model. *IEEE Trans. Information Theory*, 35(5):1044–1047, 1989.

[18] P. Piret. On the connectivity of radio networks. *IEEE Trans. Information Theory*, 37(5):1490–1492, 1991.

[19] P. Santi and D. Blough. An evaluation of connectivity in mobile wireless ad hoc networks. In *Proc. IEEE DSN 2002*, pages 89–98, 2002.

[20] P. Santi, D. Blough, and F. Vainstein. A probabilistic analysis for the range assignment problem in ad hoc networks. In *Proc. ACM MobiHoc 01*, pages 212–220, 2001.

[21] S. Singh, M. Woo, and C. Raghavendra. Power-aware routing in mobile ad hoc networks. In *Proc. ACM Mobicom 98*, pages 181–190, 1998.

[22] Y. Xu, J. Heidemann, and D. Estrin. Geography-informed energy conservation for ad hoc routing. In *Proc. ACM Mobicom 01*, pages 70–84, 2001.

APPENDIX

PROOF OF THEOREM 2. Let $d = 1$, and subdivide $[0, l]$ into $N = \frac{2l}{r}$ non-overlapping cells of side $\frac{r}{2}$. It is immediate that if every segment contains at least one node, then the resulting communication graph is connected. Let $\mu_0(n, N)$ be the random variable denoting the number of empty cells. Since $\mu_0(n, N)$ is a non-negative integer random variable, then $P(\mu_0(n, N) > 0) \leq E[\mu_0(n, N)]$, where $E[\mu_0(n, N)]$ is the expected value of $\mu_0(n, N)$ ([15], pp. 10-11). We have [13]:

$$E[\mu_0(n, N)] = N \left(1 - \frac{1}{N}\right)^n$$

We want to investigate the asymptotic value of $E[\mu_0(n, N)]$ as $l \to \infty$, which, given the hypotheses $r = r(l) \ll l$ and $n = n(l) \gg 1$, is equivalent to the asymptotic for $N, n \to \infty$. Taking the logarithm, we obtain:

$$\ln E[\mu_0(n, N)] = \ln N + n \ln \left(1 - \frac{1}{N}\right) = \ln \frac{2l}{r} + n \ln \left(1 - \frac{r}{2l}\right) \quad (3)$$

Since $r/l \to 0$ as $l \to \infty$, we can approximate the last term of (3) with the first term of its Taylor expansion, obtaining:

$$(3) \approx \ln \frac{2l}{r} - \frac{nr}{2l} \quad (4)$$

Substituting the expression $rn = kl \ln l$ in (4), we obtain:

$$(4) = \ln \frac{2l}{r} - \frac{k \ln l}{2} = \ln \frac{2}{r l^{k/2-1}}$$

If $k > 2$, or if $k = 2$ and $r = r(l) \gg 1$, then

$$\lim_{n, N \to \infty} \ln E[\mu_0(n, N)] = -\infty \ ,$$

hence $\lim_{n,N\to\infty} E[\mu_0(n, N)] = 0$ and $\lim_{l\to\infty} P(\mu_0(n, N) = 0) = 1$. It follows that each cell contains at least one node a.a.s., which implies $\lim_{l\to\infty} P_{conn}(l) = 1$.
The proof for the cases $d = 2$ and $d = 3$ are similar, and are obtained by subdividing R into non-overlapping d-dimensional cells of side $\frac{r}{2\sqrt{2}}$ and $\frac{r}{2\sqrt{3}}$, respectively. □