

Toward Quasiregular Sensor Networks: Topology Control Algorithms for Improved Energy Efficiency

Xiaowen Liu, *Student Member, IEEE*, and Martin Haenggi, *Senior Member, IEEE*

Abstract—Uniformly random or Poisson distributions are widely accepted models for the location of the nodes in wireless sensor networks if nodes are deployed in large quantities and there is little control over where they are dropped. On the other hand, by placing nodes in regular topologies, we expect benefits both in coverage and efficiency of communication. We describe and analyze a basic localized algorithm and three modifications for topology control that provide a tradeoff between performance and deployment cost. The objective is to regularize the topology for improved energy efficiency. The basic algorithm produces quasiregular networks, which only use nodes as sentries and relays that are approximately evenly spaced, thereby emulating a regular grid topology. It is shown that quasiregular networks have a significant energy and lifetime advantage compared with purely random networks. We consider two specific types of quasiregular networks: the ones that are based on a Gaussian deviation about an ideal grid point (type A), and the ones that consist of a subset of nodes taken from a Poisson point process (type B). We show that the two types are equivalent for a certain density of the Poisson point process and, in particular, that in both cases the deviation from the ideal regular grid follows a Rayleigh distribution, whereas the distance between nearest neighbors is Ricean.

Index Terms—Wireless sensor networks, wireless communications, network topology, network protocols, Poisson point processes, Rayleigh fading.

1 INTRODUCTION

IN theoretical studies and simulations, the nodes of large sensor networks are often assumed to be randomly distributed, either uniformly or as a Poisson point process, which seems to be a good model for certain modes of deployment, for example, when nodes are dropped in large numbers from an airplane. On the other hand, depending on the application, it may also be possible to place sensors at equal distances, for example, in a square grid. In so doing, we expect benefits both in coverage and efficiency of communication [1].

Uniformly random and completely regular topologies are the two extreme cases. For some applications, a model that incorporates some uncertainty into a regular distribution may be more realistic, as it may not be possible to deploy nodes completely regularly.

The idea to partition the network area into regular square grid cells has been explored for energy-saving purposes. Xu et al. [2] came up with virtual grids which are defined such that the nodes in one square cell can communicate with all the nodes in the neighboring cell. In that way, nodes in one cell are considered equivalent for routing. So, only one node needs to be active in each cell, while the other nodes can sleep to save energy. The

problems with this model are that there may be empty cells, that active nodes may still be very close, and that nodes need to be able to transmit reliably over distances larger than twice the length of the cell.

Due to the large variance in the internode distances, it is very difficult to efficiently communicate and balance the energy consumption in a network with uniformly random distribution. Hence, it is highly desirable to make the node distribution more regular by selecting an appropriate subset of random nodes. We describe and analyze a basic localized algorithm and three modifications for topology control that achieve this objective of regularizing the topology for improved energy efficiency while maintaining the coverage properties. The basic algorithm produces *quasiregular networks*, where only nodes that are approximately evenly spaced to emulate a regular grid network are active and other nodes are put to sleep to save energy. After nodes remain active for a period of time, the virtual grid is shifted and nodes closest to the shifted grid points are active for the same period of time (or phase). We analyze the network lifetime of quasiregular networks in two operating modes, a *monitoring mode* where a subset of nodes is active, acting as sentries, and a *reporting mode*, where an event of interest has been detected and a set of nodes forming a route to a base station is relaying messages. It is shown that quasiregular networks substantially outperform random networks in both modes.

Section 2 gives two definitions of quasiregular networks and introduces two types of quasiregular networks: Type A, where the coordinates of the nodes are Gaussian distributed with the mean given by regular grid points; Type B, where the selected nodes are a subset of a Poisson point process

• The authors are with the Department of Electrical Engineering, University of Notre Dame, 275 Fitzpatrick Hall, Notre Dame, IN 46556.
E-mail: {xliu4, mhaenggi}@nd.edu.

Manuscript received 15 Feb. 2005; revised 28 May 2005; accepted 25 Oct. 2005; published online 26 July 2006.

Recommended for acceptance by I. Stojmenovic, S. Olariu, and D. Simplot-Ryl.

For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDISS-0131-0205.

and every selected node is closest to a regular grid point. For networks of type B, a basic local topology control algorithm is also given. Section 3 shows that quasiregular networks of types A and B are equivalent if the density of the Poisson point process is appropriately chosen. Section 4 provides a detailed analysis of quasiregular networks of type B. It further gives a detailed analysis of the network reliability for different phases. In Section 5, it is shown that quasiregular networks of types A and B turn the Rayleigh internode distance distribution of random networks into a Ricean distribution, resulting in a significant advantage in terms of energy consumption and lifetime. To extend the network lifetime, three improved algorithms are introduced in Section 6, and their distance properties are analyzed. Section 7 concludes the paper.

2 A TOPOLOGY CONTROL ALGORITHM FOR SENSOR NETWORKS

In regular networks, the nodes are placed on the vertices of a regular grid. Here, we focus on square grids (square lattices). In (purely) random networks, the position of the nodes constitute a Poisson point process with density λ . Note that λ does not affect the *relative* distances, since all the distances are simply scaled by $1/\sqrt{\lambda}$ compared with the network with $\lambda = 1$. *Quasiregular* networks are networks that are more regular than the Poisson point process, but not perfectly regular. We offer two definitions. To be concise, we focus on infinite networks. Let R be the distance to the nearest neighbor of a node that lies within a sector $\pi/2$ of a desired direction (the source-destination axis). For a fair comparison, the network has to be normalized such that $\mathbb{E}[R] = 1$. In the Poisson case, this corresponds to a network with density 1. Both definitions are based on a measure of the uncertainty in R .

1. A *quasiregular network* is a network where the differential entropy $h(R)$ [3] (expressed in nats) satisfies $-\infty < h(R) < 1 + \frac{\gamma_{em} - \log \pi}{2}$, where γ_{em} is the Euler-Mascheroni constant. The upper bound is the differential entropy of the Rayleigh distribution with mean 1, which is the distribution of R in a Poisson point process [4].
2. A *quasiregular network* is a network with $0 < \text{Var}[R] < 4/\pi - 1$. Again, the upper bound is the variance of the Rayleigh distribution with mean 1.

The closer a network is to the lower bounds, the more regular it is. The two definitions seem equivalent in the sense that they order networks in the same way, since the relationship $h(R) \propto \text{Var}[R]$ holds for other distributions than the Rayleigh distribution [3, p. 225].

We will focus on a particular type of quasiregular networks, namely, the ones that can be obtained by *thinning* a random network. The resulting subnetwork only activates nodes as sentries and relays that are approximately evenly spaced, thereby emulating a regular topology. For example, as shown in Fig. 1, a network with uniformly randomly distributed nodes (marked by circles) can emulate a regular square network by appropriately selecting a subset of random nodes. In the first phase, the nodes closest to the integer grid points (marked by squares) are selected to be

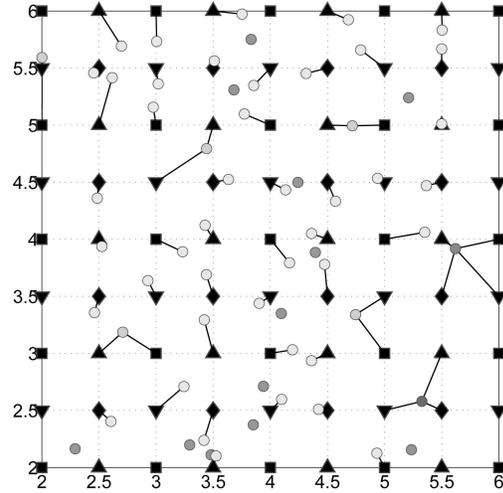


Fig. 1. A part of a network where nodes are uniformly distributed in an area 100×100 . The circles denote the random nodes. The squares, up-triangles, diamonds, and down-triangles denote the active grids in the first, second, third, and fourth phase and all of them constitute the dense grid.

active and all the other nodes are put to sleep. In the second phase, the original grid points are shifted and a new set of nearest nodes to the shifted grid points (marked by up-triangles) are selected. In each phase, the grid that the selected nodes are closest to is called the *active grid*. For example, in Fig. 1, in the first, second, third, and fourth phase, the active grid points are marked by squares, up-triangles, diamonds, down-triangles, respectively. The density of the active grid in each phase is one, without loss of generality due to scale-invariance. The grid consisting of all the original integer grid points and all the shifted grid points is denoted as *dense grid*. The active grid is always a subset of the dense grid. In this way, we can construct a quasiregular network. Next, we formally define two specific infinite quasiregular networks:

Quasiregular network of type A: Gaussian distribution. For each grid point $(x_i, y_i) \in \mathbb{Z}^2$, place a node in the plane with coordinates (X_i, Y_i) with $X_i \sim \mathcal{N}(x_i, \sigma^2)$, $Y_i \sim \mathcal{N}(y_i, \sigma^2)$, where $\sigma^2 < 1/(2\pi)$.

Quasiregular network of type B: Subset of Poisson point process. Denote the set of vertices of a Poisson point process in \mathbb{R}^2 with density $\lambda > 1$ by P . Network B consists of the smallest subset $S \subset P$ of nodes as follows: For each $p \in P \setminus S$ and any grid point, there is a node $s \in S$ such that s is closer to that grid point than p , i.e., for all $s \in S$, $\exists z_i \in \mathbb{Z}^2$ s.t. $s = \arg \min_{p \in P} \|C(p) - z_i\|$, where $C(p) \in \mathbb{R}^2$ are the coordinates of point p .

The uncertainty in the nearest-neighbor distance is reduced since the probability that it is very small or very large is substantially smaller than for a purely random network. So, both types A and B networks are indeed quasiregular.¹

This definition of quasiregular network of type B implies a *basic* local topology control algorithm to achieve quasiregularity (of type B): By exchanging position information

1. For a detailed analysis, please refer to Theorem 1, Theorem 3, and Section 5.3.

with its neighbors, each node determines whether it is closest to a virtual integer grid point.

Basic algorithm:

1. Perform synchronization and localization of the network nodes.
2. Calculate distances to the nearest grid points, exchange this information with neighboring nodes, and decide whether to enter sleep mode or stay active as a sentry.
3. After a certain period, wake up all nodes, shift the virtual grid by a certain amount. Go back to Step 2 unless the desired number of periods has passed.

Note that this is a local algorithm since it is fully distributed and only requires local data exchange. Many distributed synchronization and localization algorithms have been proposed for sensor networks, see, e.g., [5], [6] and references therein. The outcome of the localization step is that all the nodes know their position with respect to a common coordinate system, i.e., a joint grid, which is exactly what is required for Step 2 in the algorithm. The number of neighbors with which each node needs to share its distance information is limited; it does not exceed the average number of nodes within a finite radius that is of the order of the grid distance. This algorithm will be analyzed in detail in Section 4, and three modifications will be suggested in Section 6 to overcome its shortcomings.

Note that switching periods or phases incurs a substantial expenditure of energy, since all nodes need to be woken up first before Steps 2 and 3 of the basic algorithm can be carried out. Therefore, it is normally preferred to perform phase shifts only if necessary, i.e., when the currently active set of nodes is about to run out of energy. Also, phase shifts should only happen during monitoring mode. The detection of an event of interest by a sentry is assumed to cause the network to switch from monitoring to reporting mode. In reporting mode, the active set of nodes should not be changed to not perturb the ongoing transmission and avoid rerouting. If fresh nodes are being added to the network, they can be naturally integrated at the beginning of the next phase. As the phase shifts, the new nodes are considered part of the network, and after localization and synchronization, they may be selected as active nodes in the next phase.

3 PROPERTIES OF QUASIREGULAR NETWORKS

Theorem 1. *The distributions of quasiregular networks of types A and B are equivalent if $2\pi\sigma^2 < 1$ and*

$$\lambda = \frac{1}{2\pi\sigma^2}. \quad (1)$$

By equivalence, we mean that the distances between the integer grid point and its nearest neighbor node are identically distributed.

Proof. For network type A, the distance

$$D = \sqrt{(X_i - x_i)^2 + (Y_i - y_i)^2}$$

from node (X_i, Y_i) to the grid point (x_i, y_i) is Rayleigh distributed with mean $\mathbb{E}[D] = \sigma\sqrt{\pi/2}$ since the square root of the square summation of two Gaussian random variables is Rayleigh distributed. For network type B, the distance from an arbitrarily chosen point to its nearest node is also Rayleigh distributed with mean $1/(2\sqrt{\lambda})$ [4]. In particular, this is true if the arbitrarily chosen point is a grid point. So, for $\lambda = 1/(2\pi\sigma^2)$, the two distributions are identical. \square

In practice, we may consider finite areas and uniformly random distributions rather than Poisson point processes. We expect Theorem 1 to hold with good accuracy if the number of nodes is large, in which case the uniform distribution is equivalent to the Poisson process for all practical purposes.² For example, for a network of type A, consider an area $[-\frac{1}{2}, \frac{19}{2}]^2$ and place 100 nodes close to the integer square grid points with $\sigma = 1/\sqrt{2\pi} \cdot 16 \approx 0.0997$ in X and Y . This yields a grid with Gaussian uncertainty. Manual placement (with some Gaussian uncertainty) as for networks of type A can be costly and impractical, so we focus on type B, where we start with a Poisson point process and apply thinning to make it more regular. This thinning procedure is exactly the topology control algorithm described in the previous section. In the subsequent analysis, we therefore focus on networks of type B. So, for the network of type B, place N nodes uniformly randomly in the same area and pick the 100 nodes closest to the 100 active grid points. Due to the localization (Step 1 in the basic algorithm), the node can easily determine whether they are closest to an active grid point. Since the area is 100, $N = 1600$. So, for each phase, almost $1 - 100/1600 \approx 94\%$ nodes can be put to sleep. They will be activated later when the grid is shifted. For the quasiregular network to emulate a square regular network in all phases, the phase number n_p and the shift interval δ are related by $n_p = 1/\delta^2$. The shift interval and the density λ can, in principle, be chosen independently. However, the case where the number of dense grid points equals the number of points in the Poisson point process is of particular importance and will henceforth be referred to as the *natural choice*. For the previous example, if the active grid has density one, the shift interval δ should be $\delta = 1/\sqrt{\lambda}$ so that the total number of selected nodes in all the phases is approximately the total number of random nodes. In this case, the total number of phases is $n_p = 1/\delta^2 = \lambda$, which implies that for the natural choice the number of phases equals the density, i.e., $n_p = \lambda$.³ The grid shift selection scheme of the natural choice for $\lambda = 16$ is shown in Fig. 2b, where A, B, C , and D are the original grid points in the active grid for the first phase. The 16 circles within the dashed box except A are the 15 shifted grid points of the original grid point A . The shift interval is $\delta = 1/\sqrt{\lambda} = 1/4$. Since a particular node may be closest to both an original grid point and the shifted grid points, the node could be selected several times (see Fig. 1). The *usage number* U of a node in a quasiregular network of type B is

2. Note that, conditioned on the number of nodes in an area, the distribution of points in a Poisson process is uniformly random.

3. For noninteger density λ , $n_p = \lceil \text{round}(\sqrt{\lambda}) \rceil^2$, where rounding is used to obtain an integer that is close to the natural choice.

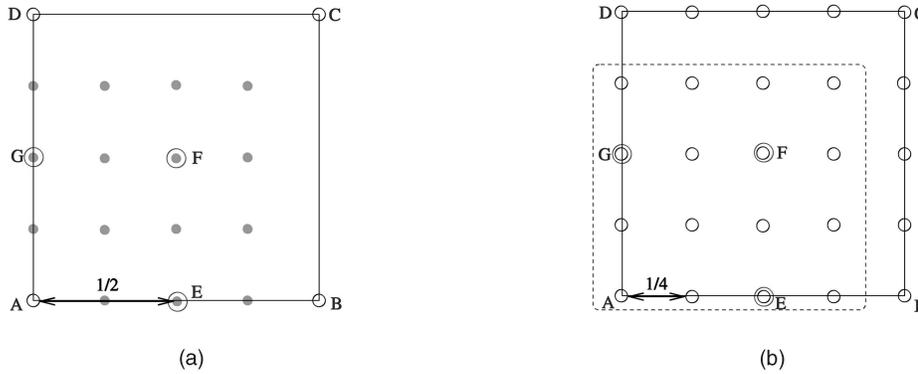


Fig. 2. Shift assignment of quasiregular network of type B. where $A, B, C,$ and D are the original square grid points with density one. (a) $\lambda = 4$, natural choice has four phases and $\delta = 1/2$, where for original grid point A (first phase), the shifted grid points are $E, F,$ and G (second, third, and fourth phases). For the 16-phase case, the 15 filled circles consist of the 15 shifted grid points of the original grid point A . (b) $\lambda = 16$, natural choice has 16 phases and $\delta = 1/4$, where the 16 circles within the dashed box except A are the 15 shifted grid points of the original grid point A . For the 4-phase case, E, F, G are the three shifted grid points of the original grid point A . (a) $\lambda = 4, n_p = 4,$ or 16 . (b) $\lambda = 16, n_p = 4,$ or 16 .

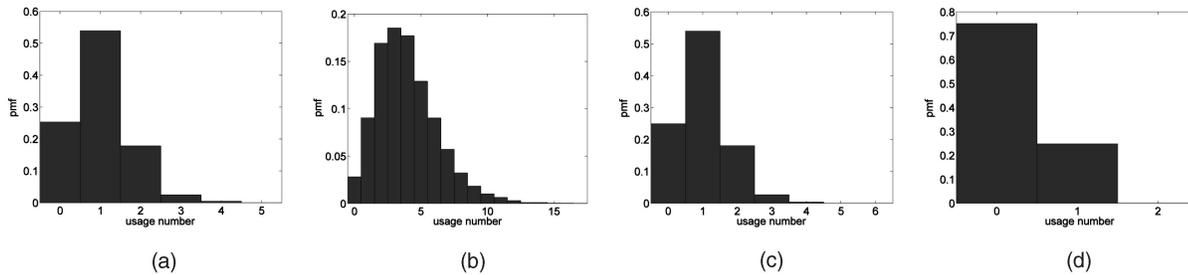


Fig. 3. Normalized histogram (pmf) of node usage numbers for $\lambda = 4, 16,$ and different phase numbers. (a) and (c) are natural choices. (a) $\lambda = 4,$ 4-phase. (b) $\lambda = 4,$ 16-phase. (c) $\lambda = 16,$ 16-phase. (d) $\lambda = 16,$ 4-phase.

defined as the number of dense grid points a node is closest to. For the natural choice, the mean usage number is one because

$$\mathbb{E}[U] = \sum_{i=0}^{\infty} i \cdot U_i = \frac{\text{density of dense grid}}{\text{density of Poisson points}} = 1, \quad (2)$$

where U_i denotes the probability mass function (pmf) of U , i.e., $U_i = \mathbb{P}[U = i]$ with $0 \leq i < \infty$, which means the probability that a node is selected by i dense grid points.

We focus on a specific network with node density $\lambda = 4$. The phase number for the natural choice is 4. The grid shift selection scheme is shown in Fig. 2a, where $A, B, C,$ and D are the original grid points in the active grid for the first phase. For the natural choice with 4 phases, for original grid point A (first phase), the shifted grid points are $E, F,$ and G (second, third, and fourth phases). The dense grid consists of the original grid points and the shifted grid points. The shift interval is $\delta = 1/\sqrt{\lambda} = 1/2$. In the next section, a detailed analysis on the node usage number U will be provided.

4 ANALYSIS OF NODE USAGE

4.1 Numerical Investigation

To determine how often a node is selected, we simulated 10^9 points of the Poisson process. For $\lambda = 4$, in addition to the natural choice, we also consider another shift value by increasing the phase number from 4 to 16. As shown in Fig. 2a, for the 16-phase case, the 15 filled circles consist of

the 15 shifted grid points of the original grid point A . The normalized histograms (probability mass functions or pmfs) of the usage numbers for the natural choice and the 16-phase case are illustrated in Figs. 3a and 3b.

For node density $\lambda = 16$, the natural choice has 16 phases. We also study the case with four phases. As shown in Fig. 2b, for the 4-phase case, $E, F,$ and G are the three shifted grid points of the original grid point A . The normalized histograms of the node usage numbers for the natural choice and the 4-phase case are shown in Fig. 3c and Fig. 3d.

An interesting observation from comparing Fig. 3a and Fig. 3c is that the normalized histograms of the node usage number with natural choice are similar, e.g., the probability that a node is not active is approximately 15 percent in Figs. 3a and 3c. Furthermore, we can see that employing more phases than the natural choice decreases the number of nodes that are not active, as expected.

4.2 Asymptotic Behavior

If the shift interval δ gets smaller and smaller, the number of phases increases. In the limiting case, there is an infinite number of shift phases so that the usage number of a node will be proportional to the area of the Voronoi cell of that node. Fig. 4a plots the normalized histogram of the usage numbers for 64 phases as $\lambda = 4$. Fig. 4b displays the normalized histogram of the Poisson Voronoi cell area (solid curve) which match the generalized gamma distribution (dashed curve):

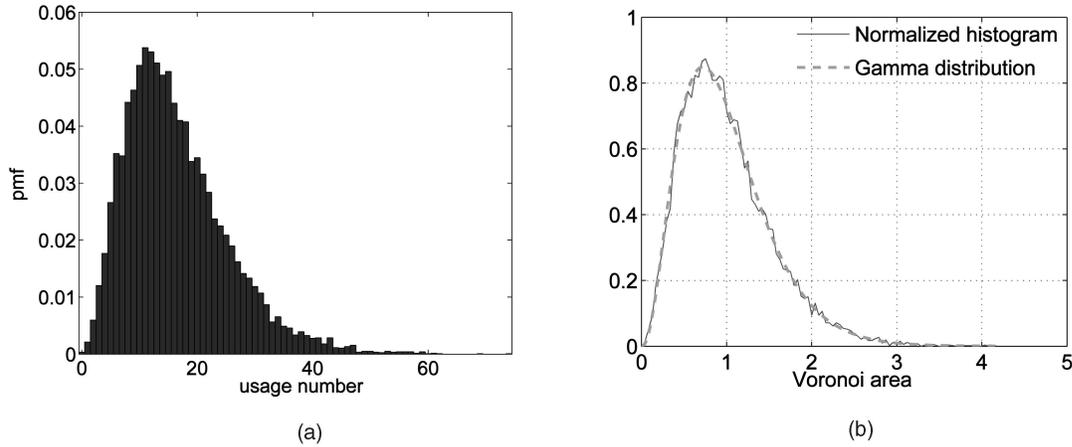


Fig. 4. (a) Normalized histogram (pmf) of usage numbers of nodes for 64 phases for $\lambda = 4$. (b) Normalized histogram (pmf) of Voronoi cells area (solid curve) and the generalized gamma distribution in (3) (dashed curve).

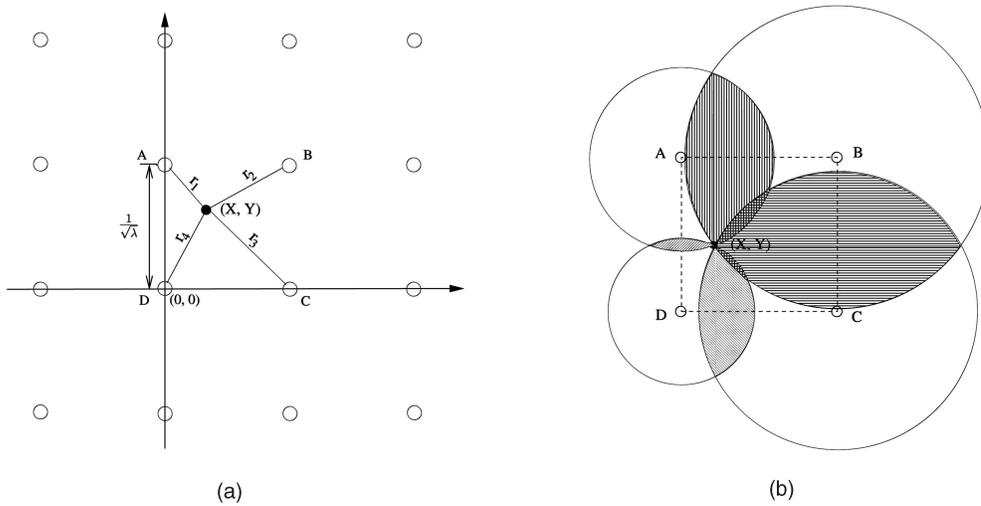


Fig. 5. In a quasiregular network, node (X, Y) and its four nearest neighbor grid points $A, B, C,$ and D . The distance from node (X, Y) to $A, B, C,$ and D is r_1, r_2, r_3, r_4 . The shift interval for the natural choice is $\delta = \frac{1}{\sqrt{\lambda}}$, so the coordinates of $A, B, C,$ and D are $(0, \frac{1}{\sqrt{\lambda}}), (\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}), (\frac{1}{\sqrt{\lambda}}, 0),$ and $(0, 0)$.

$$f(x|a, b, c) = \frac{ab^{c/a}}{\Gamma(c/a)} x^{c-1} \exp(-bx^a) \quad (a, b, c > 0), \quad (3)$$

where $a = 1.07950, b = 3.03226,$ and $c = 3.31122$ are from [7].

4.3 Analytic Bounds

The natural choice for the density of the underlying Poisson process for networks of type B is appealing since it provides a good trade-off between regularity and hardware cost. An exact and complete analysis of the usage numbers for this case is elusive. It is, however, possible to derive sharp bounds.

4.3.1 Probability That a Node is Not Active

Here, we determine the lower bound of the probability that a node is not active. The exact probability calculation is given in the Appendix. If a node is not active, it is not the nearest neighbor of any grid point. In particular, it is not the nearest neighbor of its four neighbor grid points. In addition, it is not the nearest neighbor of any more distant grid points. We consider the probability that a node at (X, Y) is not the nearest neighbor of its four neighbor grid

points A, B, C, D^4 (shown in Fig. 5a) because this is the most likely event. Note here the grid points are A and its shifted versions $B, C,$ and D . We have

$$\mathbb{P}[\text{node } (X, Y) \text{ is not the nearest neighbor of } A, B, C, \text{ and } D] \geq \hat{p} = \mathbb{E}[(1 - e^{-\lambda\pi r_1^2})(1 - e^{-\lambda\pi r_2^2})(1 - e^{-\lambda\pi r_3^2})(1 - e^{-\lambda\pi r_4^2})]. \quad (4)$$

\hat{p} is the lower bound since (4) does not consider the overlaps between the circles (the shaded areas in Fig. 5b). It is assumed that we need four different other nodes that are closer to the four cell corner grid points, although two or three nodes may be sufficient if one or more lies in the intersection of any two circles. The probability that a node is not active by its four nearest-neighbor grid points but is active by more distant grid points is rather small and neglected here. Plugging the coordinate of $A, B, C,$ and D

4. Here, the points $A, B, C,$ and D are different from the integer grid points $A, B, C,$ and D in Fig. 2.

$(0, \frac{1}{\sqrt{\lambda}}), (\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}), (\frac{1}{\sqrt{\lambda}}, 0), (0, 0)$ into (4), and considering the uniform distribution of X and Y , we obtain

$$\begin{aligned} \hat{p} &= 1 - \operatorname{erf}^2(\sqrt{\pi}) + \operatorname{erf}\left(\frac{\sqrt{2\pi}}{2}\right)\operatorname{erf}\sqrt{2\pi}e^{-\frac{\pi}{2}} + \operatorname{erf}^2\left(\frac{\sqrt{2\pi}}{2}\right)e^{-\pi} \\ &\quad - \frac{1}{3}e^{-\frac{\pi}{3}}\left[\operatorname{erf}\left(2\sqrt{\frac{\pi}{3}}\right) + \operatorname{erf}\left(\sqrt{\frac{\pi}{3}}\right)\right]^2 + \frac{1}{4}e^{-2\pi}\operatorname{erf}^2(\sqrt{\pi}) \approx 0.2362. \end{aligned} \quad (5)$$

For $\lambda = 4$, the fraction of never active nodes is 24.68 percent (see Fig. 3a), which confirms the above lower bound.

4.3.2 Probability That a Node is Active

The probability that a node is activated by its nearest grid point is

$$p = 4\lambda \int_0^{\frac{1}{2\sqrt{\lambda}}} \int_0^{\frac{1}{2\sqrt{\lambda}}} e^{-\lambda\pi(x^2+y^2)} dx dy = \operatorname{erf}\left(\frac{\sqrt{\pi}}{2}\right)^2 \approx 0.624, \quad (6)$$

which is a lower bound of the probability that a node is active.

4.3.3 Probability That a Node is Selected Four Times

As shown in Fig. 5b, the area covered by the four circles is smallest if the random node (X, Y) is in the center of the square. In this case, the area is $\frac{1}{\lambda}(2 + \pi)$, so the probability that the area has no other random nodes is

$$p = e^{-\lambda\frac{1}{\lambda}(2+\pi)} \approx 0.58\%, \quad (7)$$

which is an upper bound of the probability that a node is selected four times.

4.4 Reliability Analysis for Different Phases

In this section, we assume that the network is in monitoring mode, i.e., that most of the nodes' energy is consumed to stay awake for surveillance. So, we define the node lifetime L as the duration of a node being continuously active (awake), which is identical for every node. The lifetime of network type B is defined as the time during which in each phase at least a threshold ratio η of the selected subset is alive. So, in the following, we will determine the fraction of live nodes in each phase. Again, we first consider the important natural choice with $\lambda = 4$ for the reliability analysis. For network type B, we define the subset S_i ($i \in \mathbb{N}$), to be the set of selected nodes in the i th phase. We extend this notation by introducing S_0 —the set of nodes that are never used. The subset of nodes that are selected in different phases are not disjoint, i.e., nodes may be selected repeatedly. As shown in Fig. 1, some nodes are selected in several phases, which means they are the nearest neighbor of several dense grid points (connected to two, three, or four dense grid points by black lines). It is important to find the probability that a node belongs to multiple S_i , i.e., is selected in several phases. To this end, we introduce a *probability measure* $\mu(\cdot)$ as follows: $\mu(\cdot) : \mathcal{S} \rightarrow [0, 1]$ is the probability that a node of the original Poisson point process belongs to a set $S \in \mathcal{S}$, where \mathcal{S} is the σ -algebra of the Poisson point set⁵ and, therefore, constitutes a measurable

space with $\mu(S_0 \cup S_1 \cup S_2 \cup \dots) = 1$, so $\mu(\cdot)$ is a probability measure. With U_i the pmf of the usage number U (the fraction of nodes that are selected i times), we have $\mu(S_1 \cup S_2 \cup \dots) = 1 - U_0$. Moreover, as shown in Fig. 5b, the probabilities that a node is selected in phases 1 and/or 2 and/or 3 and/or 4 (nearest neighbor of D and/or C and/or B and/or A) have the following equalities due to the homogeneity of the Poisson point process:

$$\begin{aligned} \mu(S_1) &= \mu(S_2) = \mu(S_3) = \mu(S_4) =: \mu_1 \\ \mu(S_1 \cap S_2) &= \mu(S_2 \cap S_3) = \mu(S_3 \cap S_4) = \mu(S_4 \cap S_1) =: \mu_{2n} \\ \mu(S_1 \cap S_3) &= \mu(S_2 \cap S_4) =: \mu_{2d} \\ \mu(S_1 \cap S_2 \cap S_3) &= \mu(S_1 \cap S_2 \cap S_4) = \mu(S_1 \cap S_3 \cap S_4) \\ &= \mu(S_2 \cap S_3 \cap S_4) =: \mu_3 \\ \mu(S_1 \cap S_2 \cap S_3 \cap S_4) &=: \mu_4. \end{aligned} \quad (8)$$

So, we denote μ_1 as the probability that a node belongs to S_i with $i > 0$; μ_{2n} as the probability that a node is selected by two nearest-neighbor dense grid points, e.g., $\mu(S_1 \cap S_2)$ (selected by D and C in Fig. 5b); μ_{2d} as the probability that a node is selected by two diagonal dense grid points, e.g., $\mu(S_2 \cap S_4)$ (selected by C and A in Fig. 5b); μ_3 as the probability that a node belongs to the intersection of three of these sets, etc. Our simulation considering 10^9 points (as shown in Fig. 3a) indicates $U_0 = 24.6829$ percent, $U_1 = 54.1292$ percent, $U_2 = 18.0906$ percent, $U_3 = 2.7125$ percent, $U_4 = 0.3644$ percent, and $U_5 = 0.0189$ percent. $U_5 > 0$ shows that there is a small fraction of nodes that actually are selected by two points in the active grid in the same phase. For higher i , U_i becomes too small to be seen in the figure. Since the U_i values are very small for $i > 4$ (and exponentially decreasing), we can safely ignore them and assume that only U_0 through U_4 are nonzero. In terms of probabilities, this means that we are looking at nodes that are selected at most four times only. Analogously to μ_{2n} and μ_{2d} , there are two different probabilities that a node is selected by two dense grid points, we denote them as U_{2n} (nearest neighbor) and U_{2d} (diagonal), so $U_2 = U_{2n} + U_{2d}$. From the simulation, we obtain $U_{2n} = 16.7661$ percent, $U_{2d} = 1.3245$ percent. Note that there are five μ values and five U values and there is a one-to-one relationship between them. For example, the probability that a node is only selected in phase 1 (but not selected in phases 2, 3, or 4) can be expressed as:

$$\begin{aligned} \mu(S_1 \setminus (S_2 \cup S_3 \cup S_4)) &= U_1/4 = \mu(S_1) - \mu(S_1 \cap S_2) \\ &\quad - \mu(S_1 \cap S_4) - \mu(S_1 \cap S_3) \\ &\quad + \mu(S_1 \cap S_2 \cap S_3) + \mu(S_1 \cap S_2 \cap S_4) \\ &\quad + \mu(S_1 \cap S_3 \cap S_4) \\ &\quad - \mu(S_1 \cap S_2 \cap S_3 \cap S_4) \\ &= \mu_1 - 2\mu_{2n} - \mu_{2d} + 3\mu_3 - \mu_4, \end{aligned} \quad (9)$$

since the intersections of two and more sets have to be added and subtracted appropriately to yield the measure for $S_1 \setminus (S_2 \cup S_3 \cup S_4)$. Carrying this out for all values of U_i , the relationship between the pmf of the usage number and the measures μ_i can be summarized as follows:

5. So, in particular, \mathcal{S} includes all the possible unions and intersections of the sets S_i .

TABLE 1

Lifetime Comparison of Two Quasiregular Networks of Type B with Different Phase Number n_p and Density λ

	$\lambda = 4$		$\lambda = 9$		$\lambda = 16$		$\lambda = 25$	
Phase number n_p	4 (N)	16	9 (N)	4	16 (N)	25 (N)		
Lifetime ($\eta = 0.75$)	2L	L	4L	4L	6L	9L		
Lifetime ($\eta = 0.5$)	4L	2L	9L	4L	16L	25L		

The network lifetime is defined as the time during which in each phase at least a fraction η of the selected subset is alive. (N denotes the natural choice.)

$$\begin{bmatrix} U_1 \\ U_{2n} \\ U_{2d} \\ U_3 \\ U_4 \end{bmatrix} = \mathbf{T} \begin{bmatrix} \mu_1 \\ \mu_{2n} \\ \mu_{2d} \\ \mu_3 \\ \mu_4 \end{bmatrix} = \begin{bmatrix} 4 & -8 & -4 & 12 & -4 \\ 0 & 4 & 0 & -8 & 4 \\ 0 & 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_{2n} \\ \mu_{2d} \\ \mu_3 \\ \mu_4 \end{bmatrix}. \quad (10)$$

Since the matrix T is upper triangular, it is very easily invertible. The μ_i values are given by

$$\begin{aligned} [\mu_1 \ \mu_{2n} \ \mu_{2d} \ \mu_3 \ \mu_4]^t &= \mathbf{T}^{-1}[U_1 \ U_{2n} \ U_{2d} \ U_3 \ U_4]^t \\ &= [0.2498 \ 0.0591 \ 0.0238 \ 0.0104 \ 0.0036]^t. \end{aligned} \quad (11)$$

For the natural choice, the duration of a phase is assumed to be equal to the lifetime of the nodes. Because shifting the grid and activating a different set of nodes causes overhead, there is no reason to do this before the current set of nodes expires. In each phase, $1/\lambda = 1/4 = 25$ percent of all nodes are selected. Then, after phase 1, 25 percent of the nodes are dead. Since a fraction $\mu(S_1 \cap S_2)$ of the nodes are in $S_1 \cap S_2$, in phase 2 there are only $f_2 = 1 - \frac{\mu(S_1 \cap S_2)}{1/\lambda} = 1 - 4\mu_{2n}$ of the nodes in S_2 alive. Note that the f_i denote the fraction of nodes alive in phase i . Taking into account the fraction of nodes that have been active already in previous phases, we obtain for phases 3 and 4:

$$\begin{aligned} f_3 &= 1 - \frac{1}{1/\lambda} (\mu(S_2 \cap S_3) + \mu(S_1 \cap S_3) - \mu(S_1 \cap S_2 \cap S_3)) \\ &= 1 - 4(\mu_{2n} + \mu_{2d} - \mu_3) \\ f_4 &= 1 - 4(\mu(S_1 \cap S_4) + \mu(S_2 \cap S_4) + \mu(S_3 \cap S_4) \\ &\quad - \mu(S_1 \cap S_2 \cap S_4) - \mu(S_1 \cap S_3 \cap S_4) \\ &\quad - \mu(S_2 \cap S_3 \cap S_4) + \mu(S_1 \cap S_2 \cap S_3 \cap S_4)) \\ &= 1 - 4(2\mu_{2n} + \mu_{2d} - 3\mu_3 + \mu_4). \end{aligned} \quad (12)$$

From above analysis and (11), we have $f_1 = 1$, $f_2 = 0.7635$, $f_3 = 0.7099$, and $f_4 = 0.5422$. So, by simple inspection of f_1, f_2, f_3 , and f_4 , it is straightforward to obtain the first column of Table 1 (which is the natural choice for $\lambda = 4$): For the threshold $\eta = 0.75$, only f_1 and f_2 are greater than η , so there are two phases that have an alive node percentage greater than $\eta = 0.75$, which results in a lifetime of $2L$. We also consider the natural cases for $\lambda = 9, 16$, and 25 and include the results in Table 1. For all the natural cases, we can proceed as in (12) to obtain the f_i values. The details are omitted here. For $\lambda = 4$ and 16 , we also consider non-natural cases, see Table 1.

For the case $\lambda = 4$, $n_p = 16$, one may also decide to switch phases after $L/4$, so that the maximum duration of the entire network is $4L$, as in the natural case $\lambda = n_p = 4$. The resulting lifetime is $2.5L$ for $\eta = 0.75$ and $3.75L$ for $\eta = 0.5$. Although this is slightly better for $\eta = 0.75$ than the natural case, the energy consumption to switch phases needs to be considered, too, and is likely to offset the benefit of choosing a larger n_p . In conclusion, the natural choice best enhances the lifetime of quasiregular networks. Note that for a regular network with unity density, the lifetime is L . This analysis confirms that emulating a regular network from a random one indeed increases the network lifetime, at the price of more nodes deployed in the network. As can be seen from the four natural cases considered in Table 1, increasing the node density results in longer lifetime, so there is a tradeoff between hardware cost and lifetime.

If the node density λ is not exactly i^2 for some $i \in \mathbb{N}$, the phase number has to be chosen as $n_p = \lceil \text{round}(\sqrt{\lambda}) \rceil^2$ to ensure that each node is used approximately once, i.e., to get close to the natural case. If $n_p > \lambda$, nodes will be selected more than once on average, and if $n_p < \lambda$, some nodes will never be used. The above choice of n_p best balances these two problems of shortened lifetime and waste of nodes.

5 COMPARISON OF THE ROUTE LIFETIME FOR DIFFERENT NETWORKS

In this section, we assume the network operates in reporting mode, i.e., there is a phenomenon of interest detected in the network, causing heavy traffic along at least one route. In this case, the lifetime of this route is determined by the transmit (and possibly receive) energy consumption. We will focus on the former.

5.1 The Rayleigh Fading Link Model

We assume a narrowband Rayleigh block fading channel. A transmission from node i to node j is successful if the signal-to-interference-plus-noise ratio (SINR) γ_{ij} is above a certain threshold Θ that is determined by the communication hardware and the modulation and coding scheme [8]. The SINR γ is given by $\gamma = Q/(N_0 + I)$, where Q is the received power, which is exponentially distributed with mean \bar{Q} . Over a transmission of distance d with an attenuation factor α , we have $\bar{Q} = P_0 d^{-\alpha}$, where P_0 denotes the transmit power, N_0 the noise power, and I is the interference power affecting the transmission, i.e., the sum of the received power from all the undesired transmitters. The analysis is simplified by the following Theorem [9]:

Theorem 2. *In a network where all channels are affected by independent Rayleigh fading and nodes transmit at power level P_i ($i = 0, \dots, k$), the reception probability $\mathbb{P}[Q_0 \geq \Theta(I + N_0)]$ of a transmission over a link distance d_0 with transmit power P_0 and k other nodes at distance d_i can be factorized into the reception probability of a zero-noise network and the reception probability of a zero-interference network, i.e.,*

$$p_r = \underbrace{\exp\left(-\frac{\Theta N_0}{P_0 d_0^{-\alpha}}\right)}_{p_r^N} \cdot \underbrace{\prod_{i=1}^k \frac{1}{1 + \Theta \frac{P_i}{P_0} \left(\frac{d_0}{d_i}\right)^\alpha}}_{p_r^I}. \quad (13)$$

p_r^N is the probability that the SNR $\gamma^N := Q_0/N_0$ is above the threshold Θ , i.e., the reception probability in a zero-interference

network as it depends only on the noise. The second factor p_r^I is the reception probability in a zero-noise network.

This allows an independent analysis of noise and interference issues. From (1), we have

$$P_0 = \frac{d_0^\alpha \Theta N_0}{-\ln p_r^N}. \quad (14)$$

If the link reception probability p_r^N is fixed, we can adapt the transmit power based on (2). Note we only consider the case when p_r^N is fixed since $P_0 \propto d_0^\alpha$ gives constant p_r^N . p_r is much more difficult to keep constant because the interference is likely to change in every timeslot, in particular if the traffic is bursty. The link reception probability decreases monotonically with distance if the transmit power is fixed, and in networks with nondeterministic node placement, the link energy consumption is proportional to $\mathbb{E}[d_0^\alpha]$.

5.2 Regular and Random Networks

The lifetime of a route is determined by the maximum energy consumption among the nodes in the route. In regular networks, we assume the nodes are placed on an integer square grid over an area $[0, m] \times [0, m]$, and the next-hop receiver of each node is one of the four nearest neighbors. For random networks, the Poisson point process has density one, and the nodes are distributed in the same area. So, random networks have the same size and node densities as regular networks.

5.2.1 Random Networks

With power control: We adapt the transmit power to d^α to compensate for the path loss (see (2)) and employ the generic routing strategy from [4]: Each node in the path sends packets to its nearest neighbor that lies within a sector ϕ , i.e., within $\pm\phi/2$ of the source-destination direction. The internode distance R (node distance between nearest neighbors) along one route is Rayleigh distributed i.e., $f_R(x) = x\phi e^{-x^2\phi/2}$ [4]. The expected value of d^α is $\mathbb{E}[R^\alpha] = \left(\frac{2}{\phi}\right)^{\alpha/2} \Gamma(1 + \frac{\alpha}{2})$. The energy consumption is decreasing with increasing ϕ . The maximum energy consumption in an h -hop route

$$\mathbb{E}[R_{max}^\alpha | h] = \mathbb{E}[\max\{R_1^\alpha, R_2^\alpha, \dots, R_h^\alpha\}]$$

is given by:

$$\mathbb{E}[R_{max}^\alpha | h] = \int_0^\infty [1 - (F_{R^\alpha}(y))^h] dy, \quad (15)$$

$$F_{R^\alpha}(y) = \mathbb{P}[R^\alpha \leq y] = 1 - e^{-\frac{y^2/\alpha\phi}{2}},$$

where h is the hop number and $F_{R^\alpha}(y)$ is the cumulative distribution function (cdf) of R^α . We use a routing sector $\phi = \pi/2$ for random networks (with power control) which is equivalent to nearest-neighbor routing in regular square grid networks. As was derived in [10], for $\alpha \geq 2$

$$\begin{aligned} \mathbb{E}[\max\{R_1, R_2, \dots, R_h\}^\alpha] &\geq (\mathbb{E}[\max\{R_1^2, R_2^2, \dots, R_h^2\}])^{\frac{\alpha}{2}} \\ &> \mathbb{E}[R^2](\ln h + \gamma_{em})^{\alpha/2}, \end{aligned} \quad (16)$$

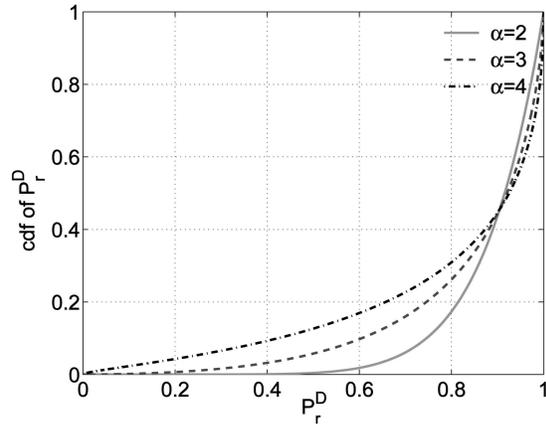


Fig. 6. Cdf of P_r^D for random networks with $\gamma_N = 10$.

where $\gamma_{em} \approx 0.5772$ is the Euler-Mascheroni constant. Thus, for $\alpha \geq 2$, the maximum energy consumption is at least logarithmically increasing with the hop number h . For example, for a 30-hop route, $\mathbb{E}[R_{max}^4 | 30]$ is about 28, indicating that the lifetime of this route is only 1/28 of that of a regular network, where $\mathbb{E}[R^\alpha] = 1$.

In a square regular network with unit density, the energy consumption is the same for all nodes in a route, i.e., the normalized energy consumption is $\mathbb{E}[R^\alpha] = \mathbb{E}[R_{max}^\alpha | h] = 1$ for any h . In a random network (with $\phi = \pi/2$ and $\alpha = 4$), however, we obtain $\mathbb{E}[R^\alpha] = 3.6$ and $\mathbb{E}[R_{max}^\alpha | h = 10] = 16.4$ for a 10-hop route and $\mathbb{E}[R_{max}^\alpha | h = 20] = 23.6$ for a 20-hop route, respectively. So, the lifetime of the routes in random networks is considerably shorter.

Without power control: Since strategies with equal transmit power are energy balanced, we also study equal transmit power schemes. Let the normalized SNR $\gamma_N := P_0/(\Theta N_0)$, from (13), resulting in a link reception probability $p_r^N = e^{-\frac{d^\alpha}{\gamma_N}}$. Considering the link distance d a random variable, then $P_r^D = e^{-\frac{d^\alpha}{\gamma_N}}$ is essentially a transformation of the random variable d and has a cdf as following:

$$\begin{aligned} F_{P_r^D}(y) &= \mathbb{P}[P_r^D \leq y] = \mathbb{P}\left[e^{-\frac{d^\alpha}{\gamma_N}} \leq y\right] = \int_{(-\gamma_N \ln y)^{\frac{1}{\alpha}}}^\infty x\phi e^{-\frac{x^2\phi}{2}} dx \\ &= e^{-\frac{(-\gamma_N \ln y)^{\frac{2}{\alpha}}\phi}{2}}. \end{aligned} \quad (17)$$

Note that for $\alpha \rightarrow \infty$, $F_{P_r^D}(y) = e^{-1}$ for $0 \leq y < 1$; and $F_{P_r^D}(y) = 1$ for $y \geq 1$. So, the limiting case is actually a disk model.

Fig. 6 illustrates the cdf of P_r^D for $\alpha = 2, 3, 4$ and $\gamma_N = 10$ in random networks. It is shown that with medium transmit power, very small link reception probabilities exist with a certain probability. If $\gamma_N = 10$ and $\alpha = 4$, for example, the probability that P_r^D is below 20 percent is 4.3 percent. Over an h -hop route, the cdf of the minimum reception probability is:

$$\begin{aligned} \mathbb{P}[\min\{P_{r1}^D, P_{r2}^D, \dots, P_{rh}^D\} \leq y] &= 1 - (\mathbb{P}[P_r^D > y])^h \\ &= 1 - (1 - F_{P_r^D}(y))^h. \end{aligned} \quad (18)$$

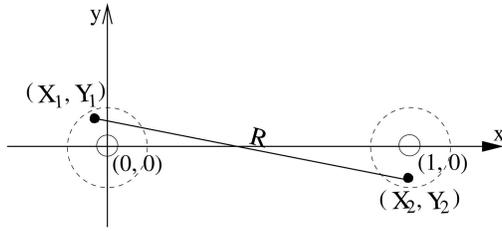


Fig. 7. Internode distance in a quasiregular network.

For example, for a 30-hop route, with $\gamma_N = 10$, corresponding to $P_r^D = 90$ percent for $d = 1$, there is a 14 percent chance that the minimum p_r^N is below 1 percent for $\alpha = 4$. This illustrates that strategies without power control will suffer from either very low end-to-end throughput or very high energy consumption (and interference) over short links.

5.3 Quasiregular Networks

By comparing the energy consumption for random and regular networks, we demonstrated that random distributions incur substantially higher energy expenditures. The large variance in the link distances necessitates power control with a large dynamic range, which, in turn, entails a proportional variance in the nodes' lifetime. The only way to avoid these fundamental problems is to abandon the principle of nearest-neighbor routing and only use nodes as relays that are approximately evenly spaced.

Theorem 3. *In quasiregular networks (of type A or B), the node distance between nearest neighbors (internode distance) follows a Ricean distribution.*

Proof. The distances in the x and y -axes of a node to the ideal grid point are Gaussian random variables. As shown in Fig. 7, we assume the ideal grid points are $(0, 0)$ and $(0, 1)$, the real location of the two nodes is (X_1, Y_1) and (X_2, Y_2) . Thus, $X_1 \sim \mathcal{N}(0, \sigma^2)$, $X_2 \sim \mathcal{N}(1, \sigma^2)$, and $Y_1 \sim \mathcal{N}(0, \sigma^2)$, $Y_2 \sim \mathcal{N}(0, \sigma^2)$. We have

$$\begin{aligned} \Delta X &= X_2 - X_1 \sim \mathcal{N}(1, 2\sigma^2), \\ \Delta Y &= Y_2 - Y_1 \sim \mathcal{N}(0, 2\sigma^2). \end{aligned} \quad (19)$$

The internode distance $R = \sqrt{\Delta X^2 + \Delta Y^2}$ is Ricean distributed with pdf

$$p_R(r) = \frac{r}{2\sigma^2} \exp\left(-\frac{1+r^2}{4\sigma^2}\right) I_0\left(\frac{r}{2\sigma^2}\right), \quad r \geq 0, \quad (20)$$

where $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind [11]. \square

Therefore, quasiregular networks turn the Rayleigh internode distance distribution of Poisson random networks into Ricean distribution. This is analogical to turning the Rayleigh fading channel into a Ricean channel by adding a strong line of sight (LOS) component.

Now, we study the differential entropy and variance of the internode distance R to see if they meet the requirement of the two definitions of quasiregular networks. Because the variance of R comes from the variances of ΔX and ΔY , under the condition that $2\pi\sigma^2 < 1$ (see Theorem 1), it is reasonable to assume that the variance is dominated by the distance along the axis that has $\mathcal{N}(1, 2\sigma^2)$ distribution. So, the variance of R can be approximated by the variance of ΔX , namely, $2\sigma^2$. Consider the 4-phase natural choice, where $\lambda = 4$ and $\sigma \approx 0.1995$ (see Theorem 1), the variance of R is $2\sigma^2 = 1/(4\pi)$, which is less than the upper bound $4/\pi - 1$ given by definition (2) of quasiregular networks. With $2\sigma^2$ as the variance of the approximated Gaussian internode distance, we have $h(R) = 1/2 \cdot \log(2\pi e 2\sigma^2)$ [3]. So, in the case $\lambda = 4$, we obtain $h(R) = 1/2 \cdot \log(2e/4) \approx 0.15$, which is less than the 0.72 upper bound given by definition (1) of quasiregular network.

Next, we will determine the lifetime benefit that results from Ricean distances rather than Rayleigh distances. The simulation results of average maximum R^α for an h -hop route $\mathbb{E}[R_{max}^\alpha | h]$ is plotted in Fig. 8 for $\alpha = 2, 3, 4$ by solid lines. The dashed line is $\mathbb{E}[R_{max}^\alpha | h]$ for random network for $\alpha = 4$. Obviously, the maximum energy consumption in a route for quasiregular networks is much smaller than that of random networks.

6 MODIFIED ALGORITHMS FOR EXTENDED LIFETIME

In the basic topology control algorithm introduced in Section 3, there is a fraction of nodes that is never used, and there are more and more nodes missing in the quasiregular topology with increasing phase numbers. To alleviate this problem, we suggest three improvements over the basic algorithm which make better use of the nodes. The

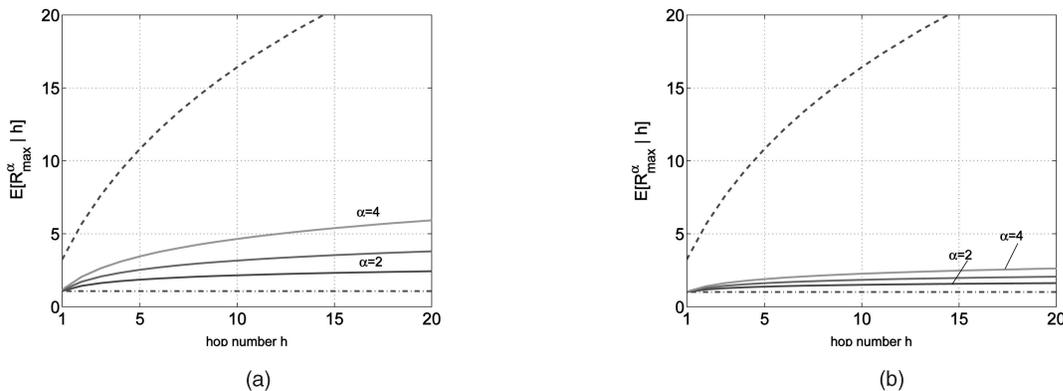


Fig. 8. $\mathbb{E}[R_{max}^\alpha | h]$ versus h for (a) $\lambda = 4$ and (b) $\lambda = 16$. The solid lines are for quasiregular networks for various α . The dashed line is for random networks for $\alpha = 4$. The dash-dotted line is for a square regular network with unit density, where $\mathbb{E}[R^\alpha] = \mathbb{E}[R_{max}^\alpha | h] = 1$. (a) $\lambda = 4$. (b) $\lambda = 16$.

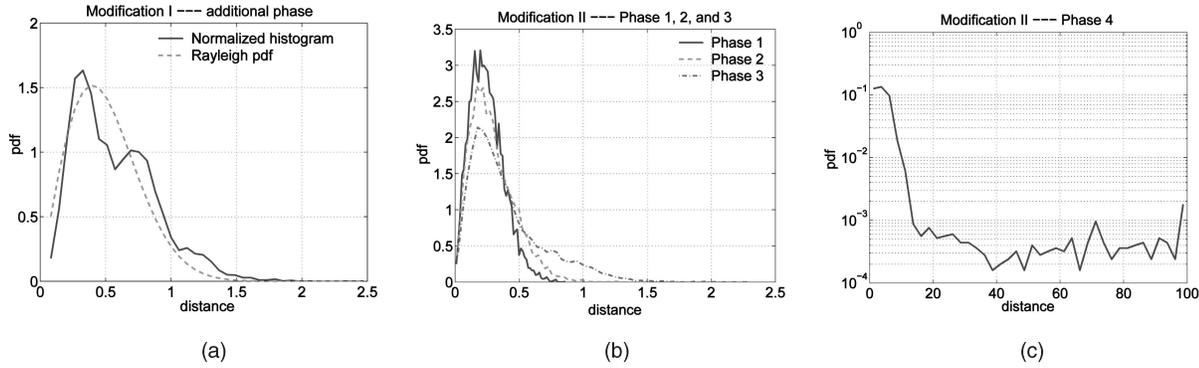


Fig. 9. For $\lambda = 4$, for Modification I, (a) comparison of the distribution of distance in the additional phase with the Rayleigh distribution with mean $1/\sqrt{\lambda} = 1/2$. For Modification II, (b) normalized distance histograms of phases 1, 2, and 3, and (c) of phase 4. Note that a logarithmic scale is used for (c) to better visualize the small but nonzero probabilities of large distances.

numerical results presented in this section are obtained from simulations with area 100×100 .

6.1 Modification I

From the previous analysis, we know that for the natural choice, approximately 25 percent (Fig. 3a and Fig. 3c) of the nodes are never activated. However, they can be turned on in an additional phase at the end. By doing so, the modified algorithm I extends the network lifetime by one more phase duration L . So, the network has *original phases* (the ones already present in the basic algorithm) and an *additional phase*. The distribution of the distance between a grid point and its nearest node in this additional phase may be approximated by the distribution of the distance between nearest neighbors in a Poisson point process with density $\lambda/4$, which is simply Rayleigh with mean $1/\sqrt{\lambda}$ [4]. The difference stems from the fact that the selected nodes in previous phases are not chosen independently. Fig. 9a displays the comparison of such two distance distributions in the additional phase for $\lambda = 4$. Note that in previous phases, the distance between a grid point and its nearest node has Rayleigh distribution with mean $1/(2\sqrt{\lambda})$. The disadvantage is that certain nodes are still selected in multiple phases.

6.2 Modification II

To avoid the problem that nodes are selected in multiple phases, Modification II lets each node be picked only once, which means once a node has been selected by one grid point, it can not be selected again even it is also the nearest node of other grid points. The advantage of Modification II is that every node is selected exactly once, which increases the lifetime in the original phases in monitoring mode compared to Modification I. The disadvantage is that the distances between the grid points and their nearest nodes grow larger at later phases (as shown in Figs. 9b and 9c). The analysis in Section 5.3 shows that larger distances from the grid points imply a higher variance in the internode distance, which results in higher energy consumption for a route in a reporting mode.

6.3 Modification III

The problem of smaller reliability in later phases and very large distance in the final phase can be solved by the third proposed modification, which is a trade-off between Modifications I and II. In each phase, pick the closest node

to the active grid point from the nodes that are alive. The simulation result ($\lambda = 4$) shows that the fraction of live nodes in Modification III for phases 2, 3, and 4 are 0.9879, 0.9195, and 0.7059, which are better than 0.7635, 0.7099, and 0.5422 of Modification I, but there are still 10.1 percent of the total number of nodes never activated. The normalized distance histograms of four phases (Fig. 10a) can be approximated by the Rayleigh distributions (Fig. 10b) with mean $1/(2\sqrt{\lambda(1-f)})$, where f is the fraction of nodes that has been selected in the previous phase(s). Note that in contrast to Modification II, the same node in Modification III may be picked by two active grid points. The advantage of Modification III is that it decreases the fraction of nodes that are selected multiple times, which increases the lifetime in the original phases in monitoring mode compared to Modification I, and it has smaller distances in higher phases than Modification II (by comparing Fig. 9c and Fig. 10a).

7 CONCLUSIONS

We proposed and analyzed topology control algorithms for improved energy efficiency. The basic algorithm turns a random network into a quasiregular network of type B, which is equivalent to the Gaussian quasiregular network of type A. It abandons the principle of nearest-neighbor routing and has every node transmit over a similar distance. This way, the nodes chosen as sentries and relays approximately form a regular subnetwork, emulating a regular topology. We suggest *differential entropy* and *variance* of the nearest-neighbor distance as measures for regularity. If the variance of the nearest-neighbor distances goes to zero, the network is completely regular. Similarly, since differential entropy is a measure for the *uncertainty* of a random variable, the higher it is for the nearest-neighbor distances, the "more random" the network topology is. The two measures are not independent. They both capture the (ir)regularity of a node distribution.

We have analyzed the network lifetime of regular, random, and quasiregular networks in two operating modes, a *monitoring mode* and a *reporting mode*. In both cases, quasiregular networks have substantial advantages over purely random ones.

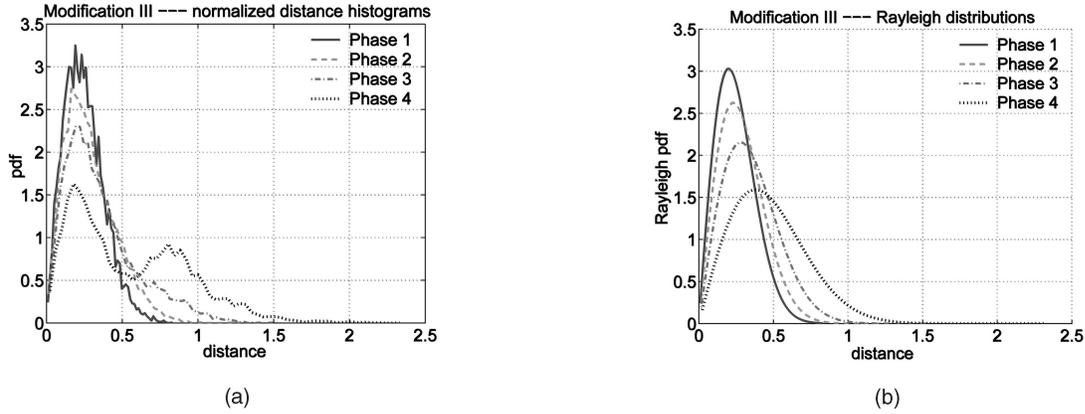


Fig. 10. For $\lambda = 4$, for Modification III, (a) normalized distance histograms of different phases and (b) Rayleigh distributions with mean 0.2500, 0.2886, 0.3523, 0.4780.

Monitoring mode: Compared with regular networks (that have λ times less nodes), the lifetime of quasiregular networks of type B is extended, i.e., the quasiregular network can to some extent exploit the higher number of nodes. For the natural choice in quasiregular networks of type B, where the number of nodes corresponds to the total number of dense grid points, approximately 25 percent of the nodes are never activated. An improved algorithm (Modification I) extends the network lifetime by activating all the unused nodes in an additional phase (but does not solve the problem that nodes may be selected repeatedly in previous phases). Modification II lets each node be selected exactly once, with the disadvantage that the distances between a grid point and its nearest live node grow large for the later phases. Modification III presents a trade-off between I and II, selecting the closest node to the active grid point from nodes that are still alive. However, there are still some nodes never activated.

Reporting mode: The comparison of the maximum node energy consumption (that determines the lifetime of the route) of a single route in networks with regular and random topology shows that regular networks drastically outperform random ones: for a path loss exponent of 4 and a route of 10-20 hops, the lifetime of a route in a random network is about $20 \times$ smaller. This analysis assumes that nodes adapt their transmit power according to the hop distance to keep the link reception probability p_r^N constant. Equal power strategies would better balance the energy consumption among the nodes but the routes would suffer from very low reception probabilities. Clearly, the cause of these problems is the variance in the transmission distances. It is shown that quasiregular networks of type B provide a solution to the energy consumption problem in random networks. Based on the premise that route longevity implies network longevity, quasiregular networks of type B outlive random networks although the number of nodes is the same.

The proposed "distance equalization" scheme also solves the problem of power amplifier inefficiencies addressed in [12], since it avoids power control over a large dynamic range and, in turn, permits the amplifiers to use operating points with high efficiency. In terms of deployment, these results suggest that one should aim at a regular node spacing whenever possible. For sensor networks, more

regular topologies also have an obvious advantage in terms of coverage [1]. By turning random into quasiregular networks by means of the proposed algorithm, we expect little or no loss in coverage, since isolated nodes will still be active. Clearly, by thinning, it is not possible to improve the coverage, but if it is done in a clever way as suggested, there is no significant loss in coverage.

APPENDIX

THE PROBABILITY THAT A NODE IS NOT ACTIVE FOR THE NATURAL CHOICE

A node is not activated means it is not closest to any dense grid points. Denote p as the probability that node (X, Y) is not the nearest neighbor of dense grid point A, B, C , and D (shown in Fig. 5). We have

$$p = 1 - p_1 + p_2 - p_3 + p_4, \quad (21)$$

where $p_1 = 4\mu_1$ is the probability that node (X, Y) is closest to any one grid point of A, B, C , and D ; $p_2 = 4\mu_{2n} + 2\mu_{2d}$ is the probability that node (X, Y) is closest to any two grid points; $p_3 = 4\mu_3$ is the probability that node (X, Y) is closest to any three grid points; and $p_4 = \mu_4$ is the probability that node (X, Y) is closest to the four grid points. We denote the area of circle centered at A as M_A , the area covered by the overlap between circles centered at A, B as $M_{A \cap B}$, the area covered by circles centered at A, B is $M_{A \cup B}$, and so on. So,

$$\begin{aligned} p_1 &= \mathbb{E}[e^{-\lambda M_A} + e^{-\lambda \pi M_B} + e^{-\lambda M_C} + e^{-\lambda \pi M_D}] \\ &= \mathbb{E}[e^{-\lambda \pi r_1^2} + e^{-\lambda \pi r_2^2} + e^{-\lambda \pi r_3^2} + e^{-\lambda \pi r_4^2}], \\ p_2 &= \mathbb{E}[e^{-\lambda M_{A \cup B}} + e^{-\lambda M_{B \cup C}} + e^{-\lambda M_{C \cup D}} + e^{-\lambda M_{D \cup A}} + e^{-\lambda M_{A \cup C}} \\ &\quad + e^{-\lambda M_{B \cup D}}], \\ p_3 &= \mathbb{E}[e^{-\lambda M_{A \cup B \cup C}} + e^{-\lambda M_{B \cup C \cup D}} + e^{-\lambda M_{C \cup D \cup A}} + e^{-\lambda M_{D \cup A \cup B}}], \\ p_4 &= \mathbb{E}[e^{-\lambda M_{A \cup B \cup C \cup D}}]. \end{aligned} \quad (22)$$

When we calculate p_3 , we neglect the area where three circles overlaps, shown by the area filled by dense square grids in Fig. 5b, since for some node locations, this area should be added, whereas in others, it should be deducted. We obtain

$$\begin{aligned}
M_{AUBUCUD} &= \pi r_1^2 + \pi r_2^2 + \pi r_3^2 + \pi r_4^2 - M_{A \cap B} - M_{B \cap C} \\
&\quad - M_{C \cap D} - M_{D \cap A} + M_{B \cap D} + M_{A \cap C}; \\
M_{AUBUC} &= \pi r_1^2 + \pi r_2^2 + \pi r_3^2 - M_{A \cap B} - M_{B \cap C} - \\
&\quad \text{sign} \left(X + Y - \frac{1}{\sqrt{\lambda}} \right) M_{A \cap C}; \\
M_{AUB} &= \pi r_1^2 + \pi r_2^2 - M_{A \cap B}.
\end{aligned} \tag{23}$$

The area covered by circles centered at A , B , and C is the sum of the area of circles centered at A , B , and C minus the overlap area of $M_{A \cap B} + M_{B \cap C}$ and minus $\text{sign} \left(X + Y - \frac{1}{\sqrt{\lambda}} \right) M_{A \cap C}$. The terms of $\text{sign} \left(X + Y - \frac{1}{\sqrt{\lambda}} \right) M_{A \cap C}$ comes from the fact that if the node is in the left-lower triangle, the area of $M_{A \cap C}$ should be added, if the node is in the right-upper triangle, the area of $M_{A \cap C}$ should be subtracted.

$$\begin{aligned}
M_{A \cap B} &= r_1^2 \arcsin \frac{1/\sqrt{\lambda} - Y}{r_1} + r_2^2 \arcsin \frac{1/\sqrt{\lambda} - Y}{r_2} \\
&\quad - \left(\frac{1}{\sqrt{\lambda}} - Y \right) \frac{1}{\sqrt{\lambda}}, \\
M_{A \cap C} &= r_1^2 \arcsin \frac{Y'}{r_1} + r_3^2 \arcsin \frac{Y'}{r_3} - Y' \frac{\sqrt{2}}{\sqrt{\lambda}}, \\
Y' &= r_1 \left| \sin \left(\frac{\pi}{4} - \arctan \frac{X}{\frac{1}{\sqrt{\lambda}} - Y} \right) \right|, \\
M_{B \cap D} &= r_2^2 \arcsin \frac{Y''}{r_2} + r_4^2 \arcsin \frac{Y''}{r_4} - Y'' \frac{\sqrt{2}}{\sqrt{\lambda}}, \\
Y'' &= r_4 \left| \sin \left(\arctan \frac{Y}{X} - \frac{\pi}{4} \right) \right|.
\end{aligned} \tag{24}$$

Note although $M_{A \cap C}$ is different from $M_{B \cap D}$ for a specific point, their average values after integration are the same. Using MATLAB's `dblquad` function, p is 0.2479 which does not depend on λ . The simulation shows that the probability that a node is not active is 0.2468 (Fig. 3a).

ACKNOWLEDGMENTS

The support of the US National Science Foundation (grants ECS 03-29766 and CAREER CNS 04-47869) is gratefully acknowledged.

REFERENCES

- [1] S. Megerian, F. Koushanfar, G. Qu, G. Veltri, and M. Potkonjak, "Exposure in Wireless Sensor Networks: Theory and Practical Solutions," *Wireless Networks*, pp. 443-454, Aug. 2002.
- [2] Y. Xu, J. Heidemann, and D. Estrin, "Geography-Informed Energy Conservation for Ad Hoc Routing," *Proc. ACM/IEEE Int'l Conf. Mobile Computing and Networking*, pp. 70-84, July 2001.
- [3] T.M. Cover and J.A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons, Inc., 1991.
- [4] M. Haenggi, "On Routing in Random Rayleigh Fading Networks," *IEEE Trans. Wireless Comm.*, <http://www.nd.edu/mhaenggi/routing.pdf>, July 2005.
- [5] B. Sundararaman, U. Buy, and A.D. Kshemkalyani, "Clock Synchronization for Wireless Sensor Networks: A Survey," *Ad Hoc Networks*, vol. 3, no. 3, pp. 281-324, May 2005.
- [6] K. Langendoen and N. Reijers, "Distributed Localization in Wireless Sensor Networks: A Quantitative Comparison," *Computer Networks*, vol. 43, no. 4, pp. 499-518, 2003.

- [7] M. Tanemura, "Statistical Distributions of Poisson Voronoi Cells in Two and Three Dimensions," *Forma*, vol. 18, no. 4, pp. 221-247, 2003.
- [8] A. Ephremides, "Energy Concerns in Wireless Networks," *IEEE Wireless Comm.*, vol. 9, no. 4, pp. 48-59, Aug. 2002.
- [9] M. Haenggi, "Analysis and Design of Diversity Schemes for Ad Hoc Wireless Networks," *IEEE J. Selected Areas in Comm.*, vol. 23, no. 1, pp. 19-27, Jan. 2005.
- [10] M. Haenggi, "Twelve Reasons Not to Route over Many Short Hops," *Proc. IEEE Vehicular Technology Conf. (VTC '04 Fall)*, Sept. 2004.
- [11] J.G. Proakis, *Digital Communications*, third ed., McGraw-Hill, Inc., 1998.
- [12] M. Haenggi, "The Impact of Power Amplifier Characteristics on Routing in Random Wireless Networks," *Proc. IEEE Global Comm. Conf. (GLOBECOM '03)*, Dec. 2003.



Xiaowen Liu received the MSc degree in signal and information processing from the Institute of Acoustics, Chinese Academy of Sciences, in 1998. She entered the University of Notre Dame as a graduate student in January 2001. She received the Master's degree in electrical engineering in 2002. Currently, she is working toward the PhD degree in the Department of Electrical Engineering. Her current research interests include wireless network and communications, especially the performance analysis of wireless ad hoc and sensor networks. She is a student member of the IEEE.



Martin Haenggi received the Dipl. Ing. (MSc) degree in electrical engineering from the Swiss Federal Institute of Technology in Zurich (ETHZ) in 1995. In 1995, he joined the Signal and Information Processing Laboratory at ETHZ as a teaching and research assistant. In 1996, he received the Dipl. NDS ETH (postdiploma) degree in information technology and, in 1999, he completed his PhD thesis on the analysis, design, and optimization of cellular neural networks. After a postdoctoral year at the Electronics Research Laboratory at the University of California in Berkeley, he has been with the Electrical Engineering Department at the University of Notre Dame since 2001, where he is currently an associate professor. For both his MSc and PhD theses, he was awarded the ETH medal, and he received a CAREER Award from the US National Science Foundation in 2005. He is a member of the editorial board of the *Elsevier Journal on Ad Hoc Networks*. His scientific interests include networking and wireless communications, with an emphasis on ad hoc and sensor networks. He is a senior member of the IEEE and the IEEE Computer Society.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.