Energy-efficient forwarding for ad hoc and sensor networks in the presence of fading

Michele Zorzi† and Ramesh R. Rao‡

†DEI – Università di Padova – via Gradenigo, 6/a – 35131 Padova, Italy – zorzi@dei.unipd.it
‡Dept. of ECE, University of California at San Diego, La Jolla, CA 92093-0407, USA – rrao@ucsd.edu

Abstract—In this paper we study the multihop performance of two energy efficient forwarding schemes (GeRaF and GAF) in a Rayleigh fading propagation scenario. Specifically, we evaluate the number of hops which are necessary to reach a destination at distance \( D \), as a function of the density of available relay nodes. Analytical and simulation results show that GeRaF significantly outperforms GAF from the multihop point of view. GeRaF’s energy/latency performance is also found to be very robust with respect to propagation impairments.

I. INTRODUCTION AND SYSTEM MODEL

Ad hoc and sensor networks are recently attracting a lot of interest in the networking research community. Many technical challenges need to be solved in this area, including such diverse topics as routing, security, MAC and application-layer as well as system-level issues. One of the key problems which need to be solved before sensor networks become widespread is energy efficiency. In fact, one cannot expect to be able to replace the batteries of all sensor nodes in a network which potentially has a broad geographical extent and a large number of nodes. Since the amount of energy which can be stored in relatively small-sized batteries is not large either, the only solution is to try and use it very judiciously, designing energy aware protocols which make the best possible use of the limited energy available.

Since the only way to really save energy in a node is to put it to sleep, a number of schemes have been recently proposed in which sleep modes are used aggressively. Of special interest in this paper is schemes in which a dense node deployment can be used to allow nodes to sleep most of the time, such as SPAN [1] and GAF [2]. The main idea behind these schemes is that one only needs to keep awake at any given time just enough nodes to provide connectivity and bandwidth.

A similar concept is the basis for Geographic Random Forwarding (GeRaF), proposed in [3], [4]. Unlike in SPAN and GAF, in which the sleep activity of the various nodes needs to be coordinated, here nodes are allowed to go to sleep and to wake up at random. In the presence of enough nodes deployed, and if the duty cycle of the sleep activity is properly chosen, it is likely that a packet will be able to find its way through the network to its final destination, using as relays the nodes which happen to be awake at the time their service is requested.

In this paper, we focus on GeRaF and GAF. Previous analyses of these two protocols as presented in [2], [3] do not take into account fading effects (which are typical in wireless environments), but rather rely on the assumption that propagation can be described in terms of coverage circles. The main contributions of this paper are the development of an analytical approach and a performance study of the multihop behavior of GeRaF and GAF in a wireless scenario characterized by Rayleigh fading and inverse power path loss. In particular, we evaluate how many hops are needed for a packet to reach a destination at a given distance from the source, which will depend on how relays are chosen.

We consider a wireless network in which nodes are randomly distributed throughout the service area according to a Poisson process with density \( \rho \) nodes per unit area. Due to the use of sleep modes, each node is available as a relay with probability \( d \), so that the density of actually available nodes is \( M' = d \rho \) and the average number of available relays within the unit circle is \( M = \pi M' \). We assume that propagation is characterized in terms of Rayleigh fading. More specifically, we assume that the received power at distance \( r \) from the transmitter is proportional to \( \alpha r^{-\eta} \), where \( \eta \) is the propagation exponent (with typical values between 2 and 4), whereas \( \alpha \) is an exponentially distributed r.v. with unit mean. The probability that a transmission across a distance \( r \) is successful can then be expressed as

\[
P_s(r) = P[\alpha r^{-\eta} > b]
\]

where \( b \) is a properly chosen threshold which depends on the noise level at the receiver and on the quality objective. In addition, we assume that the fading process is slow enough so that the value of \( \alpha \) does not change throughout a packet exchange. The above model can also be applied to the non-fading case in which we have \( \alpha = 1 \), so that in this case \( P_s(r) = 1 \) for \( r < b^{-\frac{1}{\eta}} \) and \( P_s = 0 \) otherwise. If we normalize all distances to the coverage radius for the non-fading case, then \( b = 1 \). This choice makes it possible to easily compare the performance in the presence and absence of fading for similar conditions.

II. MULTIHOP PERFORMANCE OF GE RAF

The basic idea in GeRaF is the following. We assume that nodes know their own location. Once a node has a packet to send, it sends it using some type of broadcast address while trying to act as a relay, based on how close they are to the destination. All active (listening) nodes in the coverage area will receive this packet and will assess their own priority in trying to act as a relay, based on how close they are to the destination.

As an example, consider the situation depicted in Figure 1. The source node sends out a broadcast message, and all nodes within range hear it (left part of the figure). GeRaF chooses as the relaying node the one closest to the destination (the black node in the figure). This node becomes in turn the sender, and

1Shadowing could be considered as well, but it is not expected to yield much additional insight and is not included here in order to keep the analysis simple.
hop after hop the packet will reach the final destination (the right part of the figure shows the sequence of hops for this example). We are interested in the statistics of the number of hops as a function of the various system parameters and assumptions.

Similar to the study in the absence of fading [3], we assume here a simple scenario in which nodes are uniformly distributed through the service area, according to a Poisson process in two dimensions. However, the spatial node density is in this case modulated by the probability of success of the transmission. In fact, if we assume that channel conditions do not vary for a sufficiently long time (long enough for a data exchange to take place), a physically present relay is actually usable only with probability \( P_s(r) \) which depends on its distance from the transmitter, whereas with probability \( 1 - P_s(r) \) that relay, although physically present, is unable to successfully exchange messages, and is therefore unavailable.  

\footnote{This corresponds to considering a nonuniform density for the relays, which decays as we move away from the transmitter.}

We assume that we can always choose the best relay, i.e., the one which is closest to the final destination. In this case, the relay to be chosen is not constrained to be within a given circle, although clearly relays very close from the transmitter are very unlikely to be available.

In order to first test the performance of the scheme in the presence of fading, we set up a simple simulation. We start out with a transmitter which is at distance \( D \) from the final destination. This transmitter chooses the best relay towards the destination. With probability \( \exp(-bD^n) \), the final destination can be reached in a single hop. Otherwise, potential relays are distributed uniformly in a circle of radius \( R_e \) around the transmitter, where \( R_e \) is chosen such that \( P_s(R_e) \ll 1 \) so that any relay outside this circle can be ignored. For each relay within this circle, we flip a coin with probability \( P_s(r) \) to check whether or not the relay is available, and finally, among those who have survived, we select the one closest to the destination. In the results shown, we choose \( b = 1 \), as previously discussed.

This simple simulation is run in order to evaluate the number \( n \) of hops necessary to reach the destination as a function of the distance \( D \) and of the density of active nodes \( M' \). Unlike in the case without fading, in which it must be \( n \geq D \) and \( \lim_{M' \to \infty} n = \lceil D + 1 \rceil \), in this case there is no lower bound (other than \( n \geq 1 \)) since there is always a small positive probability that the final destination is reached in a single hop. Furthermore, as the node density tends to infinity, we can find available relays arbitrarily close to the destination, and therefore the number of hops is equal to 1 with probability \( \exp(-bD^n) \), while it is equal to 2 with probability \( 1 - \exp(-bD^n) \).

The observed behavior is shown in Figures 2 and 3, where the average number of hops (with bars denoting the standard deviation) is plotted versus the node density for both the fading and non-fading cases. The value in the abscissa, which
characterizes the node density, is the average number of available relays in the unit circle, \( M \). For the non-fading case, \( M \) is the actual average number of neighbors of each node, whereas for the fading case \( M \) is to be considered as a normalized version of the node density, while the average number of nodes which can hear a transmission ("neighbors") can be found as

\[
M_\infty = \int_0^\infty M'P_s(r)2\pi r dr = \int_0^\infty M'e^{-br^\eta}2\pi r dr
\]

\[
= \frac{2M\Gamma(2/\eta)}{\eta b^{2/\eta}} = \frac{M\sqrt{\pi}}{2\sqrt{b}}
\]  

(2)

where \( \Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt \), and the last equality corresponds to the common value \( \eta = 4 \). For \( b = 1 \), the average number of nodes which can hear a transmission is equal to 0.886\( M \) for \( \eta = 4 \), to 0.903\( M \) for \( \eta = 3 \), and to \( M \) for \( \eta = 2 \), which means that in the fading case the number of neighbors is slightly less than in the case without fading (and is equal for the limiting case \( \eta = 2 \)). This is somewhat compensated by the fact that in the presence of fading some neighbors may be found beyond the non-fading coverage radius \( r = 1 \). The plots in fact show that the average number of hops to the destination is very close in the two cases.

We can provide an analytical performance study following an approach similar to the one in [3]. Consider the situation in Figure 4. The event that the advancement towards the destination is \( \zeta \leq a \) corresponds to the fact that there are no available relays in the circle of radius \( D - a \) centered in the destination. Then the considered area can be spanned using arcs of radius \( r \) centered in the transmitter, for \( r \in [a, 2D - a] \). The length of the arc of radius \( r \) is \( 2\pi r \), whereas from geometrical considerations we obtain that

\[(D - a)^2 = D^2 + r^2 - 2rD \cos \theta\]

(3)

which leads to

\[\theta(D, a, r) = \arccos \left( \frac{D^2 + r^2 - (D - a)^2}{2rD} \right)\]

(4)

for \( a \leq r \leq 2D - a \). Then the average number of available relays in the considered area is found as

\[M_r = \int_a^{2D-a} M'P_s(r)2\pi r dr \]

\[= 2M'\int_a^{2D-a} e^{-br^\eta} \arccos \left( \frac{D^2 + r^2 - (D - a)^2}{2rD} \right) r dr \]

(5)

The relay distribution in the two-dimensional space is modeled here as a nonhomogeneous Poisson process, for which the number of relays actually present in a given area is a Poisson random variable. The probability that \( \zeta \leq a \) is then the probability that there are no relays available in the considered area, i.e.,

\[P[\zeta \leq a] = e^{-M_r}\]

(6)

which is a function of both \( a \) and \( D \).

The average advancement towards a destination at distance \( D \) is then found as

\[E[\zeta(D)] = \int_0^D (1 - P[\zeta \leq a]) da\]

(7)

A simple and accurate approximation of the average number of hops to the destination can be found as \( D/E[\zeta(D)] \). This approximation is plotted in Figure 5 for \( D = 5, 20 \) and for \( \eta = 2, 4 \). From the Figure, we can see that the approximation is very good, and that the effect of the propagation exponent is not very significant, as it leads to a slight improvement for smaller \( \eta \). It should be noted here that the distances are normalized to the coverage radius, which is of course larger for \( \eta = 2 \) than it is for \( \eta = 4 \), if the same transmit power is used. On the other hand, if the transmit power is chosen in the two cases in such a way as to provide similar coverage, the performance is as shown in Figure 5.

More precisely, one can provide true bounds for the multihop performance. As demonstrated in [3], the average number of hops to reach a destination at distance \( D \) can be bounded as follows

\[\frac{D - 1}{E[\zeta(D)]} + 1 \leq E[\zeta] \leq \frac{D}{E[\zeta(1)]} + 1\]

(8)
III. MULTIHOP PERFORMANCE OF GAF

We consider here GAF [2], a scheme in which energy consumption is controlled by treating clusters of nodes as a single logical entity. More precisely, the area is divided into square regions called grids, in such a way that any two nodes in two adjacent grids can communicate regardless of their actual location (i.e., the grid size is chosen to be $\sqrt{5}$ times smaller than the coverage radius so as to guarantee reliable communication even in the worst possible case, i.e., nodes in opposite corners of adjacent grids). Routing then proceeds along a sequence of grids, in which any node in a grid can serve as representative of all others in the same grid. This makes it possible to have all nodes in a grid sleep, except one, thereby providing a reduction of the consumed energy which is proportional to the number of nodes in the grid. Actual implementation of this scheme requires means to arrange this node rotation, i.e., signaling is needed to make sure that nodes properly coordinate. Details are provided in [2].

The performance of GAF in the absence of fading can be easily found, as the grid size must be taken to be $g = 1/\sqrt{5}$ times the coverage radius and therefore the average number of neighbors is $5\pi$ and the total number of hops depends on the distance $D$ as well as the one-hop advancement which in turn depends on the relative orientation between the grid structure and the source-destination direction. It of course does not make much sense to choose a grid size smaller than $1/\sqrt{5}$ in this case, as this would result in a larger number of hops without any advantage.

The situation is different in the presence of fading, in which transmission success is not guaranteed but rather is characterized by a probability which depends on the distance between transmitter and receiver. In this case, choosing a smaller grid size corresponds to more hops but also to more reliable transmissions (i.e., fewer retransmissions). It is then possible to find an optimal grid size which optimizes the total number of transmissions needed for a packet to travel a certain distance. In this section, we study this tradeoff.

Consider a single hop between two consecutive grids. The communicating nodes are each uniformly distributed within its own grid. Suppose for the moment that the grid size is unity. The node coordinates are independent of each other and uniformly distributed. More specifically, the first node has coordinates $(x_1, y_1)$ with $x_1, y_1 \sim U[0,1]$, whereas the second node has coordinates $(x_2, y_2)$ with $x_2 \sim U[1,2], y_2 \sim U[0,1]$. As a consequence, the horizontal distance $x = x_2 - x_1$ has a triangular pdf between 0 and 2, $f_x(x) = \tau(x - 1)$, whereas $y = y_2 - y_1$ has a triangular pdf between $-1$ and 1, $f_y(y) = \tau(y)$, with

$$\tau(a) = \begin{cases} 1 + a & -1 \leq a < 0 \\ 1 - a & 0 \leq a \leq 1 \\ 0 & |a| > 1 \end{cases}$$

For this pair of nodes, the probability of a successful transmission in the presence of fading is given by

$$P_s(x, y) = e^{-b(x^2+y^2)^{\eta/2}}\varrho^n$$

where we have taken into account that $x, y$ are normalized quantities, while the actual grid size is in general given by $\varrho$. This probability of success corresponds to an average number of attempts equal to $1/P_s(x, y)$, i.e., it takes on average

$$u = 1/P_s(x, y)$$

transmissions for a message to be exchanged between the two nodes. On average, then, the number of transmissions for a message exchange between two adjacent grids can be found as

$$E[u] = \int_0^2 dx \tau(x - 1) \int_{-1}^1 dy \tau(y) e^{b(x^2+y^2)^{\eta/2}}\varrho^n$$

That is, in general when making “one hop” between adjacent grids, one makes on average $E[u] - 1$ zero-hops (i.e., failed transmissions) plus one successful transmission.

Let us now look at the number of grids which need to be crossed to cover a distance $D$. For simplicity, let us again assume that the grid size is unity. Also, let grids be aligned along the axes of a two-dimensional orthogonal reference so that their corners have integer coordinates. Finally, let the source node be placed in the grid whose lower left corner is placed in $(0,0)$. The actual coordinates $(a_x, a_y)$ of the source node are independent and uniformly distributed in $[0,1]$. If $D$ is the distance between the source and the destination, the destination node has coordinates $(a_x + D \cos \theta, a_y + D \sin \theta)$, where $\theta \sim U[0, \pi/2]$ accounts for the random alignment with respect to the grids. We want to compute how many grids need to be crossed to go from the source to the destination. Only horizontally or vertically adjacent grids can communicate, whereas diagonal movements are not allowed. Note that in this case the number of horizontal transitions is equal to $n_x = [a_x + D \cos \theta]$ and similarly $n_y = [a_y + D \sin \theta]$ for vertical transitions. Since $E[n_x][\theta] = D \cos \theta$ and $E[n_y][\theta] = D \sin \theta$, we have that the average number of “true hops,” needed to reach the destination is given by

$$E[D \cos \theta + D \sin \theta] = \frac{4D}{\pi}$$

Taking into account that the grid size is actually $\varrho$, the above becomes

$$E[n_{tot}] = \frac{4D}{\pi \varrho}$$

We can then approximate the total number of hops by just multiplying $E[n_{tot}]$ by $E[u]$, since each true hop involves $u$ transmission attempts.

Figures 6 and 7 show some results for a distance $D = 10$ and for two typical values of $\eta$, namely 2 and 4. Each graph reports the multihop performance of both GeRaF and GAF, evaluated by analysis and simulation. It can be seen that the analytical approximation for GeRaF is very good. In the case of GAF, the match between analysis and simulation is even better. It is interesting to note that the best performance for GAF is achieved for $M \simeq 10$ (i.e., a grid size of about 0.56, which is about 25% larger than in the case without fading) and that this value depends weakly on the propagation exponent. It is also worth noting that GAF achieves significantly better performance for $\eta = 4$ than for $\eta = 2$, whereas the opposite is true for GeRaF; although in the latter case the performance difference is not very significant. In all cases, we see that GAF has significantly worse multihop performance than GeRaF.

IV. PRACTICAL SCHEME

The above results were about an ideal scheme which assumed that the best relay can always be selected. In practice, this is not possible as it would require perfect coordination among nodes. The simple implementable mechanism to make this selection proposed in [3] can be used in this case as well.
A node who wants to transmit a packet broadcasts a message. All active nodes in its coverage area will hear the message. Each node will then determine its own distance from the final destination and assess its own adequacy as a relay. This is done by first dividing the coverage area in two parts: the relay region, which contains all points closer to the destination than the transmitting node, and the non-relay region, which contains all other points. Nodes in the non-relay region are never selected as relays.

Further, the relay region is sliced in $N_p$ “priority regions,” based on the distance $\gamma$ from the destination. More specifically, region $A_i$ contains all points in the coverage region whose distance from the final destination is $\xi_{i-1} \leq \gamma < \xi_i$, $i = 1, 2, \ldots, N_p$. In [3], in which the potential relays must lie in the coverage circle, it was taken $\xi_i = D - i/N_p$, $i = 0, 1, \ldots, N_p$. In the present case, a relay can theoretically provide an advancement which may be arbitrarily close to $D$, or can be located beyond the destination itself, and therefore the relay regions must be defined differently. In particular, $A_i$ is a circular ring centered in the destination, with radii $\xi_{i-1}, \xi_i$.

Also, the non-uniform density of available relays calls for a non-uniform spacing of the $\xi_i$’s. A reasonable choice is one in which the $\xi_i$’s are selected such that the average number of available relays in each priority region is the same. (Note that we always have $\xi_0 = 0$ and $\xi_{N_p} = D$.) In this case, we have

$$\int_{A_i} M' P_\eta(\phi) d\phi = \int_{A_{i-1}} M' P_\eta(\phi) d\phi \quad i = 2, \ldots, N_p \quad (14)$$

Notice that the above integrals depend on $D$, and therefore the boundaries of the priority regions should be recomputed at every hop. In reality this dependence is weak and, for example, in the case $N_p = 4$, we numerically found that choosing the $\xi_i$’s as $\{0, D - 0.635, D - 0.39, D - 0.19, D\}$ provides very similar values for the average number of delays in the various priority regions for $D \geq 2$, and therefore can be taken as a practical choice. In any case, the choice of the threshold values could be optimized, even though preliminary results suggest that not much can be gained compared to the case proposed here. As an additional remark, we note that although the above values were found assuming $\eta = 4$, they are reasonably good for $\eta = 2$ as well. That is, using the same $\xi_i$’s as before for $\eta = 2$, while not leading to priority regions with the same number of available relays, gives a good approximation of this situation, which basically removes the need to optimize the $\xi_i$’s with respect to $\eta$.

Figure 8 shows that, as was the case in the absence of fading, a relatively small number of priority regions achieves a multihop performance very close to the optimum.

V. ENERGY-LATENCY PERFORMANCE OF GERaF IN THE PRESENCE OF FADING

In this section, we present some preliminary results about the energy/latency performance of GeRaF in the presence of fading.

As in [4] it is possible to develop an analytical approach which counts the energy expenditure of a node due to several tasks. The only difference here is that the relay population is not constrained to being in a “coverage circle” but can be anywhere, although with different probabilities. In particular, the average number of possible relays in the fading case has been computed in (2). Note that not all these neighbors are
useful as relays, as some of them may actually lead farther away from the destination. In fact, only those within the circle of radius $D$ centered in the destination are closer to the destination than the transmitter. Let $\xi'$ be the fraction of the $M_\infty$ neighbors that are in this “relay region.”

The analytical approach presented in [4] can still be used, but some modifications need to be made in order to take into account the differences outlined above. The analysis leads to closed-form formulas for the energy consumption. A good approximation of the normalized energy consumption is

$$\psi_0 \simeq \lambda \left( e^{\xi'M_\infty} - 1 \right)^{-1} \left( 3N_p + 1 \right) T_{SIG} + T_{DATA} + d + P_s / P$$

while the latency is

$$T_S = \left( e^{\xi'M_\infty} - 1 \right)^{-1} \left( 1 + 2N_p + 2x \right) T_{SIG}$$

where $d$ is the listening duty cycle, $\lambda$ is the packet arrival rate at each node, $N_p$ is the number of priority classes, $T_{SIG}$ is the duration of a signaling packet (RTS, CTS, ACK, or CTS response), $T_{DATA}$ is the duration of a data packet, $P_s$ and $P$ are the powers dissipated by the radio in sleep mode and active mode, respectively, and $x$ is the average length of the contention phase (see [4] for more details on these quantities).

As an example of the results obtained, Figures 9 and 10 provide the energy-latency tradeoff with and without fading, for two values of the propagation exponent ($\eta = 2$ and 4), and of the node density ($N = \pi \rho$ is the average number of physically deployed nodes in the unit circle). As $\xi'$ depends on the distance between source and destination, in the fading case we show pairs of approximate curves which correspond to upper and lower bounds for the possible value of $\xi'$. A more accurate analysis is out of the scope of this paper, but we do not expect that it would provide much different insights than seeing here. The results in these figures show the robustness of GeRaF with respect to the propagation parameters. In fact, it is clearly seen that the presence of fading does not lead to virtually any performance degradation. This is due to the fact that the performance of the protocol only depends on how many nodes are instantaneously available as relays, and this quantity is very insensitive to the considered propagation effects. On the other hand, protocols such as STEM [5], where for each transmission a specific node is addressed as the next hop, may be very vulnerable to fading. More precisely, protocols like STEM which address packets to a specific node often fail in the presence of fading (if the link is faded), whereas in GeRaF the packet is relayed by nodes which hear the transmission (and therefore are not faded).

VI. CONCLUSIONS

In this paper we have studied the multihop performance of two energy efficient forwarding schemes (GeRaF and GAF) in a Rayleigh fading propagation scenario. Specifically, the number of hops which are necessary to reach a destination at distance $D$ was evaluated, as a function of the density of available relay nodes. Analytical and simulation results showed that GeRaF significantly outperforms GAF from the multihop performance point of view. Also, GeRaF was found to be very robust with respect to propagation impairments. A practically implementable scheme was shown to provide essentially the same performance as the ideal GeRaF relay selection mechanism. Finally, some preliminary results which show the robustness of the energy/latency performance of the GeRaF scheme with respect to the presence of fading have also been presented.

Further analytical developments as well as a complete energy-latency analysis remain as future work in this area. Also, the effect of cross traffic and interference due to multiple source-sink pairs is being investigated.

REFERENCES