HYBRID-ARQ BASED INTRA-CLUSTER GEOGRAPHIC RELAYING

Matthew C. Valenti

West Virginia University

Morgantown, WV 26506-6109

mvalenti@wvu.edu

Bin Zhao

Efficient Channel Coding, Inc.

Brooklyn Heights, OH 44131

bzhao@eccincorp.com

ABSTRACT

This paper * *presents an integrated, cross-layer protocol that utilizes position location (e.g. through an onboard GPS receiver) and jointly performs the operations of physical-layer cooperative diversity, hybrid-ARQ retransmission, and relaying/routing. The protocol is given the descriptive name Hybrid ARq-Based INtra-cluster GEographically-informed Relaying (HARBINGER) and generalizes both Geographic Random Forwarding (GeRaF) and point-to-point hybrid-ARQ. A version called Slow-HARBINGER is analyzed in detail. Numerical results indicate a dramatic improvement in the energy-latency tradeoff as compared with conventional multihop and GeRaF.*

INTRODUCTION

A common technique for saving energy in wireless sensor networks is to periodically put each radio into a sleep mode, since listening to idle channels consumes considerable processing and transceiver power [1]. The lifetime of such networks is primarily a function of the duty cycle of the nodes, and networks whose nodes are in a sleep state for a higher percentage of time will last longer. Several protocols have been recently proposed for sensor networks with sleeping nodes. A complete survey is outside the scope of this paper, but the interested reader is referred to [2] and the references therein. Such protocols can behave in one of two ways: (1) Nodes awaken according to a deterministic rendezvous schedule, and (2) Nodes cycle on-and-off at random. This paper focuses on the second type of protocol, as it lends itself to simpler implementation by allowing each node to autonomously set its own sleep schedule and allows for accurate analysis through the application of probability theory.

If nodes know their own position and messages are addressed by location, rather than by MAC address, then this geographic information could guide the routing mechanism. A recent protocol that uses this concept is Geographic Random Forwarding (GeRaF) [2]. With GeRaF, the message is broadcast to all nodes within range and the one node that both decodes the message and is closest to the destination is the

one that forwards it. A contention strategy is proposed in [2] for determining (in a distributed fashion) which node should serve as the relay. GeRaF has the benefit of not requiring a route to be established prior to transmission, obviating the need for routing tables. Furthermore, it takes advantage of the spatial diversity in the network due to the presence of multiple nodes. In environments with fading and interference, this distributed spatial diversity could allow dramatic improvements in performance if properly exploited, i.e. by implementing *cooperative diversity* [3].

While GeRaF is effective, it tends to require a dense distribution of active nodes. If no node is within range of the source, GeRaF waits until the sleep state changes and then starts over again in the hopes that a node within range has awakened. If the density of active nodes is insufficient, the source may need to retransmit several times before there is any forward progress. However, there may be active nodes just out of the source's range that could be used. While these nodes are too far away to successfully decode the initial source transmission, they might be able to decode after the second (or later) transmission if they combine all the information they have received. This is the underlying concept behind *hybrid-ARQ* [4, 5]. With (type-II) hybrid-ARQ, each node will combine all received transmissions prior to decoding a particular message. Information can be combined in one of two ways: (1) The message is repeated by the source (repetition-coding); the receiver *diversity-combines* the repeated transmissions, (2) The source encodes the message by a low rate code and through rate-compatible puncturing a distinct portion of the codeword is transmitted each time (incremental-redundancy); the receiver *code-combines* the received code fragments [4]. With diversity-combining the receiver sees a channel with a higher effective SNR, while with code-combining it receives a code with a lower effective rate. Because the capacity of code-combining is always at least as good as the capacity of diversity-combining [4], that will be the focus of the remainder of this discussion.

By using hybrid-ARQ, performance can be improved by allowing nodes just outside of the source's range to keep previously transmitted packets to be combined with retransmissions. In an AWGN channel and in the absence of interference, all nodes within radius R_1 can be reached during

the initial transmission. If there are no geographically advantaged nodes (i.e. nodes that provide positive forward progress) within this range, the source must transmit again. GeRaF assumes that nodes outside range R_1 don't hear the message and thus the nodes that can be reached during the second transmission are limited to those only within range R_1. However, with our proposed protocol, nodes beyond of range R_1 (say, out to range R_M) may receive and maintain each transmitted packet. Thus during the second transmission the *effective* range of the source increases from R_1 to R_2. The effective range continues to increase after each retransmission until a maximum range R_M is reached. As the range increases, so does the probability of finding a geographically advantaged node. With this protocol it is possible to use a lower density of nodes than with GeRaF, yet achieve the same delay and consume almost the same (transmit) energy.

Because our protocol uses a combination of GeRaF and hybrid-ARQ, we give it the descriptive name *Hybrid ARq-Based Intra-cluster GEographically-informed Relaying (HARBINGER)*. This paper complements a companion paper [6] in describing two versions of the HARBINGER protocol with considerably different behavior. The two versions of the protocol differ primarily in the relation between the periodicity of the sleep cycle and the data packet transmission rate. In [6], we investigate a version we term Fast-HARBINGER where the sleep states are synchronized with the data packet rate. Thus, each time a message is retransmitted, the topology changes. In this paper, we investigate a version we term Slow-HARBINGER where nodes cycle in and out of sleep states at a rate that is slower than the data packet rate, thereby fixing the topology for several ARQ retransmissions (up to M).

GEOGRAPHIC RANDOM FORWARDING

Consider a two-dimensional ad hoc wireless network with nodes randomly distributed according to a Poisson process. The density of active nodes is ρ, which is much smaller than the density of all nodes (both awake and asleep). Time is slotted into intervals of τ seconds, where τ is called the *network coherence time*. During the ith network coherence interval (NCI) $t : \{(i-1)\tau \leq t < i\tau\}$, the topology remains fixed. At the end of the NCI, nodes will randomly and autonomously cycle on and off. As in [2], we assume each node knows its own position and has a circular coverage area. Nodes within the coverage circle successfully decode the initial transmission, while those outside the circle do not. This model assumes an AWGN channel with exponential path-loss and capacity-approaching channel coding (e.g. turbo or LDPC codes) but neglects the influence of interference and fading. For a discussion of the impact of block fading, the reader is referred to [7].

Once a node has a message to transmit, it will send a request-to-send (RTS) packet at the beginning of the next NCI to detect if there is a potential relay nearby. A node is said to be *geographically advantaged* if it is closer to the destination than the source is, and only geographically advantaged nodes may serve as a relay. If there is no such relay to be found, the source will send another RTS packet in the next NCI. Hopefully, through random node activity, a potential relay will appear and respond with a clear-to-send (CTS) packet indicating that it is ready to receive the subsequent data packet. If multiple potential relays respond with CTS packets, the source will use a contention scheme to choose a particular relay (ideally, the most geographically advantaged).

HARBINGER

With GeRaF if the active node density is fairly low (due to low duty cycle nodes), it is highly probably that the source will need to attempt transmission of the same message in the next NCI, which greatly increases delay. To overcome this drawback, HARBINGER incorporates hybrid Automatic Repeat reQuest (ARQ) into GeRaF. With hybrid-ARQ, distant nodes outside of the source's first attempt transmission range accumulate additional information after each retransmission until they are eventually able to decode. Equivalently, the coverage circle increases after each transmission (since the effective code rate decreases). In HARBINGER, the source will first encode the message with a low rate mother code. The mother codeword is then partitioned into M data packets, where each packet is a distinct portion of the low rate mother code (achieved through rate compatible puncturing [5]). If the network coherence time τ is long enough such that several ARQ retransmissions, i.e. M, can be attempted before the topology changes, then each message could be successfully delivered within a single NCI. This is the underlying assumption for Slow-HARBINGER. Alternatively, Fast-HARBINGER [6] intentionally changes the network topology prior to each data packet transmission to achieve an extra time diversity benefit. Notice that GeRaF is a special case of HARBINGER with $M = 1$.

HARBINGER preserves the basic packet structure, handshaking, and contention schemes in GeRaF. With Slow-HARBINGER up to M packets can be transmitted each NCI, and after the mth packet is transmitted the effective coverage area is expanded to R_m. Slow-HARBINGER has two versions (A and B). Slow-HARBINGER A maximizes message progress within each NCI by picking the relaying node that is closest to the destination. This strategy minimizes message delay. Slow-HARBINGER B minimizes the number of packet transmissions per NCI by picking a node with forward progress that is reachable with the minimum number of ARQ transmissions. This strategy minimizes transmit energy consumption, since the exponential path loss model favors many short hops over a few long ones.

A particular node that decides to receive packets will

keep every packet received so that old information may be combined with fresh information gained after each new ARQ transmission. Eventually this node will be able to decode the message, although it is possible that some other node decodes it first. If multiple nodes successfully decode the message after the same packet transmission, then a contention scheme similar to that in GeRaF could be used to choose the single relay that is most geographically advantaged. Once a specific relay is chosen, all the active nodes within the coverage area will flush their old memory (discard previously received packets) for the new message. While message flushing results in a simple protocol and tractable analysis, more sophisticated protocols can be developed that don't flush their memory and therefore experience a *cooperative diversity* effect [3, 7].

MATHEMATICAL FRAMEWORK

Let the effective coverage radius after the mth transmission be R_m. Under the assumption of capacity-approaching channel coding, exponential path loss, and AWGN, this radius can be found by first finding the channel capacity after the mth transmission:

$$C_m = \frac{m}{2} \log_2 \left(1 + K_0 d_m^{-\mu} \frac{\mathcal{E}_s}{N_o} \right) \qquad (1)$$

where $\frac{\mathcal{E}_s}{N_o}$ is the transmit signal to noise ratio, μ the path loss coefficient, K_0 the signal propagation coefficient, d_m the propagation distance. The expression (1) follows from the well known result that the capacity of parallel Gaussian channels adds [8]. Any node within the circle of radius d_m is guaranteed to correctly receive the source message of rate $r < C_m$ with no more than m transmissions, where r is the rate of each packet (assumed to be constant for all packets). Solving for the distance d_m we get

$$d_m < \left(\frac{K_0 \mathcal{E}_s/N_o}{2^{2r/m} - 1} \right)^{1/\mu} \qquad (2)$$

The radius is the maximum transmission range, i.e. $R_m = d_m$. For analytical convenience, we assume $R_1 = 1$ and normalize $\{R_m\}$ with respect to R_1, resulting in

$$R_m = \left(\frac{2^{2r} - 1}{2^{2r/m} - 1} \right)^{1/\mu} \qquad (3)$$

We consider the same coordinate system as [2, 6] with the source located at $(D, 0)$ and the destination located at $(0, 0)$. In HARBINGER, once a relay node is chosen to forward the message, it will become the source during the next NCI. Therefore, the distance between source and destination changes each time the message is successfully decoded. The hybrid-ARQ scheme results in M concentric circles centered at $(D, 0)$. Each circle O_m of radius R_m corresponds to the region that can be reached through at most m transmitted packets. We also consider a virtual circle O_0 of radius

$R_0 = 0$ for notational consistency. Likewise, we further define $D\nu$ concentric circles centered at $(0, 0)$ with each circle Q_i of radius i/ν . The concentric circles $\{Q_i\}$ quantize the whole range of possible distances from source to destination, e.g. $0 \sim D$, into $D\nu$ intervals with ν being the number of quantization interval per unit distance. The lth interval $\triangle_l = (\frac{l}{\nu}, \frac{l-1}{\nu}]$ corresponds to a region $Q_l - Q_{l-1}$.

Suppose the source is $D = \frac{j}{\nu}$ away from the destination, then partition ξ_m could be defined as

$$\xi_m = (O_m - O_{m-1}) \cap Q_j \quad \text{for } 1 \le m \le M \qquad (4)$$

Any active node in ξ_m could correctly decode the message by receiving exactly m data packets from the source. It is straightforward to show that

$$\cup_{m=1}^M \xi_m = O_M \cap Q_j \qquad (5)$$
$$\xi_i \cap \xi_k = \phi \quad \text{for } i \ne k \qquad (6)$$

Given $j - R_m\nu + 1 \le l \le j - R_{m-1}\nu$, ξ_p is further divided into three disjoint regions by $\triangle_l$ for $\forall p \ge m$. In particular,

$$\xi_p = \xi_{p,-} \cup \xi_{p,0} \cup \xi_{p,+} \qquad (7)$$
$$\xi_{p,+} = (O_p - O_{p-1}) \cap (Q_j - Q_l) \qquad (8)$$
$$\xi_{p,0} = (O_p - O_{p-1}) \cap (Q_l - Q_{l-1}) \qquad (9)$$
$$\xi_{p,-} = (O_p - O_{p-1}) \cap Q_{l-1} \qquad (10)$$

However, when $\forall p < m$, ξ_p is not further partitioned. An example with $M = 2$ is shown in Fig. 1 of [6].

PERFORMANCE ANALYSIS

In this section, we characterize the message delay and energy efficiency of both GeRaF and Slow-HARBINGER under the same mathematical framework. As we assume that data packets consume much more energy than signaling packets, the energy analysis is equivalent to finding the average number of data packet transmissions per message. Let X_i denote the event that partition ξ_i contains at least one potential relay, while $\bar{X}_i$ denotes the event that partition ξ_i contains no potential relay. Likewise, we use $X_{i,\bullet}$ to denote the event that subpartition $\xi_{i,\bullet}$ contains at least one potential relay, while $\bar{X}_{i,\bullet}$ denotes the event that subpartition $\xi_{i,\bullet}$ contains no potential relay where $\bullet$ could be '+', '0' or '-'.

GERAF

For the specific case of $M = 1$, HARBINGER reduces to GeRaF. When the source is within the coverage circle of the final destination, i.e. $D = \frac{j}{\nu}, j = 1, \ldots, \nu$, the message delay is equal to the NCI duration. On the other hand, when integer $j > \nu$, the event that the message progresses from location $D = j/\nu$ to interval $\triangle_{j-k+1}$ is equivalent to the event that there is at least one potential relay in the partition $\xi_{1,0}$ and no potential relay in partition $\xi_{1,-}$. We use progress
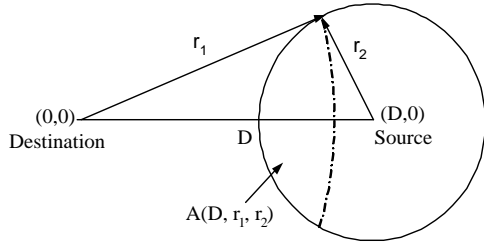
Figure 1: The intersection area of two circles of radius r_1 and r_2 separated by a distance of D.

probability $\omega(j,k)$ to denote the joint probability,

$$
\begin{aligned}
\omega(j,k) &= Pr\left\{X_{1,0} \cap \bar{X}_{1,-}\right\} \quad \text{For } k = 1, \ldots, \nu \\
&= \exp\left\{-\rho A\left(\frac{j}{\nu}, \frac{j-k}{\nu}, R_1\right)\right\} \\
&\quad - \exp\left\{-\rho A\left(\frac{j}{\nu}, \frac{j-k+1}{\nu}, R_1\right)\right\} \quad (11)
\end{aligned}
$$

where

$$
A(D, r_1, r_2) = 2\int_{D-r_2}^{r_1} \arccos\left(\frac{x^2 + D^2 - r_2^2}{2Dx}\right) x\,dx \quad (12)
$$

denotes the area of the intersection of two circles of radius r_1 and r_2 separated by a center-to-center distance of D as shown in Fig. 1. In addition, ω_0 denotes the probability such that partition ξ_1 contains no potential relay,

$$
\begin{aligned}
\omega_0(j) &= Pr\left\{\bar{X}_1\right\} \\
&= \exp\left\{-\rho A\left(\frac{j}{\nu}, \frac{j}{\nu}, R_1\right)\right\} \quad (13)
\end{aligned}
$$

We use the same recursive approach in [2] to calculate upper and lower bounds on average message delay. Because delay is an integer multiple of the NCI duration , for the remainder of the paper we normalize the delay with respect to τ. Accordingly, the upper bound $n_1(j)$ and lower bound $n_2(j)$ on delay become

$$
\begin{aligned}
n_1(j) &= 1 + \omega_0(j)n_1(j) + \sum_{k=1}^{R_1\nu} \omega(j,k)n_1(j-k+1) \\
n_2(j) &= 1 + \omega_0(j)n_2(j) + \sum_{k=1}^{R_1\nu} \omega(j,k)n_2(j-k)
\end{aligned}
$$

with initial condition $n_1(j) = n_2(j) = 1$ for $j = 1, \ldots, \nu$.

GeRaF will not transmit if there is no potential relay in the coverage area. Therefore, upper and lower bounds on the average number of data packet transmissions per message can be recursively calculated by slightly modifying the above expressions. More specifically, when integer $j > \nu$, the upper $e_1(j)$ and lower $e_2(j)$ bounds on the number of transmitted

packets become

$$
\begin{aligned}
e_1(j) &= \omega_0(j)e_1(j) + \sum_{k=1}^{R_1\nu} \omega(j,k)(e_1(j-k+1) + 1) \\
e_2(j) &= \omega_0(j)e_2(j) + \sum_{k=1}^{R_1\nu} \omega(j,k)(e_2(j-k) + 1)
\end{aligned}
$$

with initial condition $e_1(j) = e_2(j) = 1$ for $j = 1, \ldots, \nu$.

SLOW-HARBINGER A

Slow-HARBINGER A will pick the relay node closest to the destination. Consider a simple example of $M = 2$. When integer $j > R_2\nu$, the destination is outside coverage area O_2. The corresponding probability of message progress $\omega(j,k)$ is

$$
\omega(j,k) = \begin{cases} Pr\left\{\left(\cap_{i=1}^2 \bar{X}_{i,-}\right) \cap \left(\cup_{i=1}^2 X_{i,0}\right)\right\}, \\ \qquad \text{for } k = 1, \ldots, \nu \\ Pr\left\{\bar{X}_{2,+} \cap X_{2,0}\right\}, \\ \qquad \text{for } k = \nu+1, \ldots, R_2\nu \end{cases}
$$

where $\omega(j,k)$ could be further decomposed into

$$
\omega(j,k) = \omega(j,k,1) + \omega(j,k,2) \quad (14)
$$

where $\omega(j,k,m)$ denotes the joint probability of message progress from location j/ν to $\triangle_{j-k+1}$ with exactly m data packet transmissions. In particular,

$$
\omega(j,k,1) = \begin{cases} Pr\left\{\left(\cap_{i=1}^2 \bar{X}_{i,-}\right) \cap X_{1,0}\right\}, \\ \qquad \text{for } k = 1, \ldots, \nu \\ 0, \text{for } k = \nu+1, \ldots, R_2\nu \end{cases}
$$

$$
\omega(j,k,2) = \begin{cases} Pr\left\{\left(\cap_{i=1}^2 \bar{X}_{i,-}\right) \cap \bar{X}_{1,0} \cap X_{2,0}\right\}, \\ \qquad \text{for } k = 1, \ldots, \nu \\ Pr\left\{\bar{X}_{2,-} \cap X_{2,0}\right\}, \\ \qquad \text{for } k = \nu+1, \ldots, R_2\nu \end{cases}
$$

In addition $w_0(j) = Pr\left\{\cap_{i=1}^2 \bar{X}_i\right\}$.

Now consider $M > 1$. When $j > R_M\nu$, the destination is outside coverage area O_M. When $k = R_{p-1}\nu + 1, \ldots, R_p\nu$, the corresponding message progress probability $\omega(j,k)$ is

$$
\begin{aligned}
\omega(j,k) &= Pr\left\{\left(\cap_{i=p}^M \bar{X}_{i,-}\right) \cap \left(\cup_{i=p}^M X_{i,0}\right)\right\} \\
&= \exp\left\{-\rho A\left(j/\nu, (j-k)/\nu, R_M\right)\right\} \\
&\quad - \exp\left\{-\rho A\left(j/\nu, (j-k+1)/\nu, R_M\right)\right\} \quad (15)
\end{aligned}
$$

where $\omega(j,k)$ could be further decomposed into

$$
\omega(j,k) = \sum_{m=p}^M \omega(j,k,m) \quad (16)
$$

$$
\omega(j,k,m) = \begin{cases} Pr\left\{\cap_{i=p}^M \bar{X}_{i,-}\right\} Pr\left\{\cup_{i=p}^m X_{i,0} - \cup_{i=p}^{m-1} X_{i,0}\right\}, \\ \qquad \text{for } m = p, \ldots, M \\ 0, \qquad \text{for } m < p \end{cases}
$$

and

$$Pr\left\{\cup_{i=p}^{l} X_{i,0}\right\} = 1 - \exp\left\{-\rho A\left(j/\nu, (j-k+1)/\nu, R_l\right)\right\}$$
$$\exp\left\{\rho A\left(j/\nu, (j-k)/\nu, R_l\right)\right\} \quad (17)$$
$$Pr\left\{\left(\cap_{i=p}^{M} \bar{X}_{i,3}\right)\right\} = \exp\left\{-\rho A\left(j/\nu, (j-k)/\nu, R_M\right)\right\} \quad (18)$$

In addition,

$$w_0(j) = Pr\left\{\cap_{i=1}^{M} \bar{X}_i\right\} = \exp\left\{-\rho A\left(j/\nu, j/\nu, R_M\right)\right\}$$

The upper and lower bounds of average delay are

$$n_1(j) = \sum_{k=1}^{R_M \nu} \omega(j,k)(n_1(j-k+1)+1) + \omega_0(j)(n_1(j)+1)$$
$$n_2(j) = \sum_{k=1}^{R_M \nu} \omega(j,k)(n_2(j-k)+1) + \omega_0(j)(n_2(j)+1)$$

with $n_1(j) = n_2(j) = 1$ when $j = 1, \ldots, R_M \nu$. The average number of data packet transmissions is

$$e_1(j) = \sum_{k=1}^{R_M \nu} \sum_{m=1}^{M} \omega(j,k,1,m)(e_1(j-k+1)+m)$$
$$+ \omega_0(j)e_1(j)$$
$$e_2(j) = \sum_{k=1}^{R_M \nu} \sum_{m=1}^{M} \omega(j,k,1,m)(e_2(j-k)+m) + \omega_0(j)e_2(j)$$

with $e_1(j) = e_2(j) = m$ when $j = R_{m-1}\nu + 1, \ldots, R_m \nu$.

SLOW-HARBINGER B

Unlike Slow-HARBINGER A, Slow-HARBINGER B intends to minimize the data packet transmissions per NCI, thus it will increase the message delay. In particular, Slow-HARBINGER B picks the relay node that is reachable with a minimum number of ARQ retransmissions. Consider $M = 2$. The progress probability $\omega(j,k)$ is as given by (14) where

$$\omega(j,k,1) = \begin{cases} Pr\left\{\bar{X}_{1,-} \cap X_{1,0}\right\}, \\ \quad \text{for } k = 1, \ldots, \nu \\ 0, \quad \text{for } k = \nu+1, \ldots, R_2\nu \end{cases}$$

$$\omega(j,k,2) = Pr\left\{\bar{X}_1 \cap \bar{X}_{2,-} \cap X_{2,0}\right\}, \text{for } k = 1, \ldots, R_2\nu \quad (19)$$

In addition $w_0(j) = Pr\left\{\cap_{i=1}^{2} \bar{X}_i\right\}$.

Now consider $M > 1$. When $j > R_M \nu$, the destination is outside coverage area O_M. When $k = R_{p-1}\nu + 1, \ldots, R_p\nu$, its corresponding probability $\omega(j,k)$ is found as

$$\omega(j,k) = \sum_{m=1}^{M} \omega(j,k,m) \quad (20)$$

where

$$\omega(j,k,m) = \begin{cases} Pr\left\{X_{m,0} \cap \bar{X}_{m,-} \cap \left(\cap_{i=1}^{m-1} \bar{X}_i\right)\right\}, \\ \quad \text{for } m = p, \ldots, M \\ 0, \quad \text{for } m < p \end{cases}$$

with

$$Pr\left\{X_{m,0} \cap \bar{X}_{m,-}\right\} = \exp\left\{-\rho A\left(\frac{j}{\nu}, \frac{j-k}{\nu}, R_m\right)\right\}$$
$$\exp\left\{\rho A\left(\frac{j}{\nu}, \frac{j-k}{\nu}, R_{m-1}\right)\right\}$$
$$- \exp\left\{-\rho A\left(\frac{j}{\nu}, \frac{j-k+1}{\nu}, R_m\right)\right\}$$
$$\exp\left\{\rho A\left(\frac{j}{\nu}, \frac{j-k+1}{\nu}, R_{m-1}\right)\right\}$$

and

$$Pr\left\{\cap_{i=1}^{m-1} \bar{X}_i\right\} = \exp\left\{-\rho A\left(j/\nu, j/\nu, R_{m-1}\right)\right\} \quad (21)$$

On the other hand, when $j = R_{p-1}\nu + 1, \ldots, R_p\nu$, the destination is within partition ξ_p, thus is reachable with at most p data packets. In particular, $w_0(j) = 0$, and

$$\omega(j,k,m) = \begin{cases} Pr\left\{\cap_{i=1}^{p-1} \bar{X}_i\right\} \delta(m,p)\delta(j,k), \\ \quad k = R_{p-1}\nu+1, \ldots, j \\ Pr\left\{X_{m,0} \cap \bar{X}_{m,-} \cap \left(\cap_{i=1}^{m-1} \bar{X}_i\right)\right\}, \\ \quad k = R_{l-1}\nu+1, \ldots, R_l\nu, m = l, \ldots, p-1 \\ 0, \quad \text{otherwise} \end{cases}$$

Correspondingly, the upper and lower bounds of average delay as well as average data packets could be calculated recursively with initial condition $n_1(j) = n_2(j) = 1$ when $j = 1, \ldots, \nu$, $e_1(j) = e_2(j) = 1$ when $j = 1, \ldots, \nu$.

NUMERICAL RESULTS

The upper and lower bounds of message delay in HARBINGER are plotted in Fig. 2 and Fig. 3 for Slow-HARBINGER A, and Slow-HARBINGER B respectively. The bounds are calculated for $M = 2, 12$. The delay performance of GeRaF is also included ($M = 1$) as comparison. In both figures, the message delay is normalized by the corresponding network (topology) coherence time τ. The block code rate is $r = 1$ and the number of intervals per unit distance is $\nu = 50$. If we change the code rate r, the radius $\{R_i\}$ of coverage circles change accordingly which will affect the delay performance. As observed in both figures, the upper and lower bounds are quite close to each other indicating the tightness of both bounds. As the interval $1/\nu$ becomes smaller, the bounds will become more accurate and vice versa. In Fig 2, we observe that Slow-HARBINGER A significantly reduces the message delay as M increases. The result is rather intuitive. With hybrid ARQ the source/destination separation is effectively reduced by a factor of R_M due to progressive coverage expansion. It further increases the population of active nodes within the coverage region. In particular, as active node density $\rho \to \infty$, the message delay will asymptotically converge to $\lfloor \frac{D}{R_M} + 1 \rfloor$, where D is the source/destination separation. Note in Fig. 3, Slow-HARBINGER B has different delay characteristics.
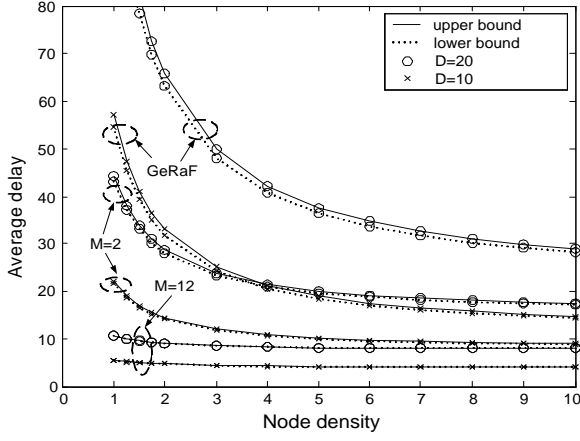
Figure 2: The average delay (normalized by τ) of Slow-HARBINGER A under different rate constraints M.
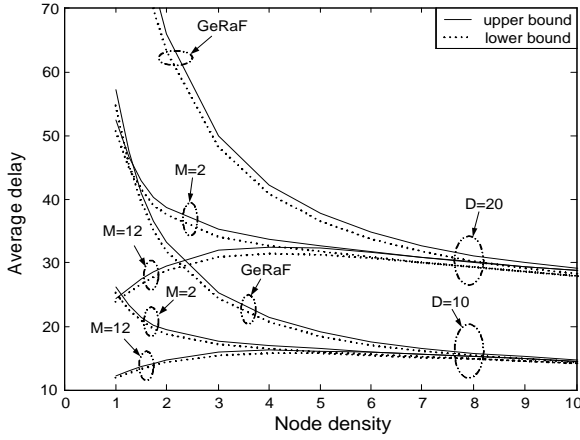

Figure 3: The average delay (normalized by τ) of Slow-HARBINGER B under different rate constraints M.

In particular, in a dense network, both GeRaF and Slow-HARBINGER B will asymptotically converge to a message delay of $\lfloor D + 1 \rfloor$ as node density $\rho \to \infty$. The performance difference between Slow-HARBINGER A and Slow-HARBINGER B in dense networks is primarily due to their different relay selection criterion. More specifically, Slow-HARBINGER A picks the relay closest to the destination while Slow-HARBINGER B picks the relay requiring minimum ARQ retransmissions.

An interesting phenomenon we observed in Fig.3 is that as the rate constraint gets fairly large, i.e. $M = 12$, the delay performance is not a monotonically decreasing function of node density. In particular, at low density, the message delay actually decreases with the node density. This observation is counter-intuitive. To explain this phenomenon, we calculate and plot the average message progress $Avg(j)$ per NCI under different node density in Fig. 4, where

$$Avg(j) = \sum_{k=1}^{R_M \nu} \omega(j,k) \frac{k}{\nu} \qquad (22)$$

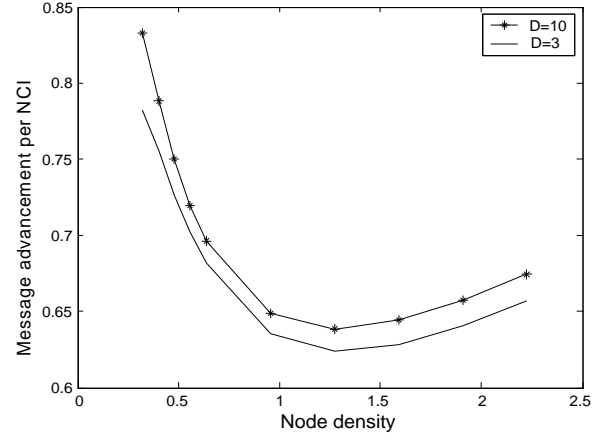

Figure 4: The average message progress per NCI for Slow-HARBINGER B under different source/destination separation.

$\omega(j,k)$ as a function of node density could be calculated through (20). Notice that in Fig. 4 the message progress is actually larger in networks with lower density, indicating that nodes closer to the destination are more likely to be chosen as relay in low density network. Therefore, its corresponding message delay becomes smaller as shown in Fig. 3.

The delay performance of HARBINGER indicates that with hybrid-ARQ, nodes are allowed to remain in a sleep state for a relatively longer percentage of time than GeRaF (for the same total node density) while still able to achieve the same delay performance as GeRaF. Alternative, with the same duty cycle of network devices, HARBINGER could significantly reduce the message delay.

The energy efficiency analysis of HARBINGER is highly dependent on the ratio of energy consumed by signalling packets to the energy consumed by data packets. As mentioned earlier, if we assume that data packets take up a majority of energy dissipation and ideally ignore the energy dissipation of signalling packets, the energy dissipation is proportional to the average number of data packet transmissions per message. Due to the proximity of both bounds, we only plot the lower bound of data packet transmissions per message for Slow-HARBINGER and Fast-HARBINGER in Fig. 5 and Fig. 6. We observe that in both figures, GeRaF actually has the best energy efficiency. HARBINGER consumes more energy than GeRaF primarily due to its relatively aggressive packet transmission strategy, non-linear coverage expansion behavior and more importantly, memory flushing mechanism. Further notice that unlike the message delay which decreases as the rate constraint increases, the energy consumption of HARBINGER actually increases along with M. In particular, with Slow-HARBINGER A the energy dissipation increases significantly as M increases in both dense and sparse networks. Unlike Slow-HARBINGER A, although the energy efficiency of Slow-HARBINGER B is
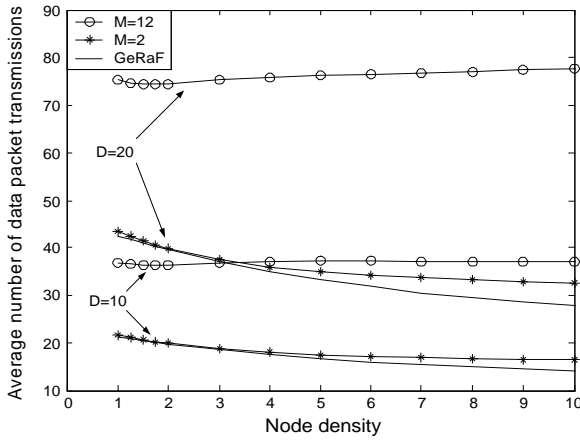
Figure 5: The average data packet transmissions per message in Slow-HARBINGER A under different rate constraints M, $M = 1$ corresponds to GeRaF.

Figure 6: The average data packet transmissions per message in Slow-HARBINGER B under different rate constraints M, $M = 1$ corresponds to GeRaF.

worse than GeRaF in low density networks, it converges to GeRaF in high density networks. In fact, as $\rho \to \infty$, Slow-HARBINGER B asymptotically requires $\lfloor D+1 \rfloor$ data packet transmissions for each message.

Altogether we investigated two different network setups where the source and destination are separated by distances of 10 and 20. We observe that both message delay and average data packets per message are almost linearly proportional to the separation distance. Further note that the above bounds are derived under the assumption of memory flushing (after each successful message transmission) which significantly reduces the relaying gain in the protocol. Therefore, in practice, without memory flushing, HARBINGER should perform much better in the sense of both message delay and energy efficiency (see [7]). Finally, we need to point out that each version of Slow-HARBINGER could be most suitable for different sensor network applications. Increasing the rate constraint does not necessarily improve the performance of HARBINGER. Rather HARBINGER with a small rate constraint, i.e. $M = 2, 3$, is appropriate for network implementation, since under small rate constraints HARBINGER could dramatically decrease the message delay without a significant increase in the energy dissipation.

CONCLUSIONS

HARBINGER is an effective cross-layer protocol for ad hoc networks that combines Geographic Random Forwarding with hybrid-ARQ. The analysis presented in this paper generalizes GeRaF, which corresponds to the specific case that $M = 1$. HARBINGER is especially beneficial over GeRaF in lower density networks when Hybrid-ARQ is applied with small rate constraint, indicating that a smaller duty-cycle sleep schedule could be used for network devices with HARBINGER, thereby increasing the useful lifetime of sensor networks. Alternatively, for the same sleep schedule,
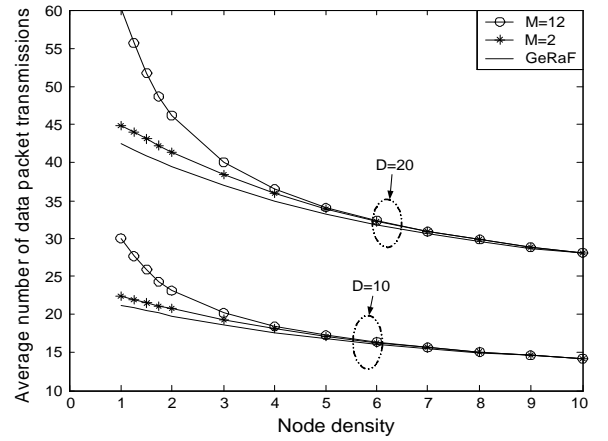
HARBINGER allows reduced end-to-end delay compared to GeRaF.

REFERENCES

[1] R. Min et al., "Energy-centric enabling technologies for wireless sensor networks," *IEEE Wireless Communications*, Aug. 2002.

[2] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance," *IEEE Trans. Mobile Comp.*, vol. 2, pp. 337–348, Oct. 2003.

[3] J. N. Laneman, *Cooperative diversity in wireless networks: Algorithms and architectures*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, Aug. 2002.

[4] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Info. Theory*, vol. vol. IT-47, pp. pp. 1971 – 1988, July, 2001.

[5] S. Wicker, *Error Control Systems for Digital Communications and Storage*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1995.

[6] B. Zhao, R. Iyer Seshadri, and M. C. Valenti, "Geographic random forwarding with hybrid-ARQ for ad hoc networks with rapid sleep cycles," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, 2004. to appear.

[7] B. Zhao and M. C. Valenti, "Practical relay networks: A generalization of hybrid-ARQ," *IEEE J. Select. Areas Commun.*, vol. 23, Jan. 2005. to appear.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.