

Geographic Random Forwarding with Hybrid-ARQ for Ad Hoc Networks with Rapid Sleep Cycles

Bin Zhao

Efficient Channel Coding
600 Safeguard Plaza, Suite 100
Brooklyn Heights, OH 44131
email: {bzhao}@csee.wvu.edu

Rohit Iyer Seshadri and Matthew C. Valenti

Lane Dept. of Comp. Sci. & Elect. Eng.
West Virginia University
Morgantown, WV 26506-6109
email: {iyerr,mvalenti}@csee.wvu.edu

Abstract— This paper proposes and analyzes a new cross-layer protocol for ad hoc and sensor networks that unifies the concepts of Geographic Random Forwarding (GeRaF) and hybrid-ARQ. The protocol is given the descriptive name *Hybrid ARQ-Based Intra-cluster GEographically-informed Relaying (HARBINGER)*. Like GeRaF, HARBINGER assumes that each node knows its own position and that messages are addressed by location. As is common for sensor networks, the nodes cycle on-and-off according to a sleep schedule. Unlike GeRaF, which returns to the initial transmission state if no active node is within range, the nodes in HARBINGER combine transmissions thereby achieving an additional time-diversity benefit. With HARBINGER, a lower density of active nodes will achieve almost the same delay and energy efficiency as GeRaF, implying that a lower duty cycle sleep schedule could be used to prolong the useful lifetime of the network. The paper gives detailed analysis of a version of the protocol (Fast-HARBINGER) whereby the sleep states of the network are synchronized with the data packet transmission rate of the network.

I. INTRODUCTION

Because of the limited energy supply of typical wireless devices, energy-efficiency is a crucial issue in the design of protocols for ad hoc networking. This is especially true for sensor networks that typically must last for a year or more without battery replacement. Perhaps the most effective way to conserve energy is to periodically put each radio into a sleep mode, since listening to idle channels consumes significant processing and transceiver power [1]. The lifetime of such networks is 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. There are two distinct categories of such protocols: (1) Nodes are awoken according to a deterministic rendezvous schedule, and (2) Nodes cycle on-an-off at random. This paper focuses on the second type of protocol, as it lends itself to simple implementation by allowing each node to autonomously set its own sleep schedule and brings to bear the analytical tools of probability theory.

If nodes know their own position and messages are addressed by location, then it is possible to use geographic

information to guide the routing mechanism. A recent protocol that uses this concept is *Geographic Random Forwarding (GeRaF)* [2]. With GeRaF, each message is broadcast to all nodes within range and the one node that both decodes the message and is closest to the destination forwards it. This has the benefit of not requiring a route to be established prior to transmission and 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 [3]. However, a key issue with GeRaF is that the relaying node must be selected in a distributed fashion, without being guided by a genie. This implies that the routing and MAC protocol must be designed jointly. A key contribution of [2] is the development of a practical MAC protocol that allows the most geographically advantaged relay to be chosen after the source has transmitted.

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 (type-II) *hybrid-ARQ* [4]. With hybrid-ARQ, packets are combined in one of two ways: (1) The receiver *diversity-combines* repeated transmissions, (2) The source encodes the message with 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. 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.

In this paper, we propose a protocol that uses a combination of GeRaF and hybrid-ARQ. We give the protocol the descriptive name *Hybrid ARQ-Based Intra-cluster*

GEographically-informed Relaying (HARBINGER). In our protocol, each prospective relay will code-combine up to M ARQ transmissions from the source. The benefit of the protocol is that it provides a better energy-latency trade-off than GeRaF, especially at lower node densities. Alternatively, the same performance can be achieved with a lower active node density, thereby prolonging the lifetime of the network. The analysis generalizes that of GeRaF, and in fact, GeRaF is a special case of HARBINGER whereby $M = 1$.

This paper complements a companion paper [5] 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 [5], 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. Thus, the topology remains fixed for several consecutive packet transmissions, i.e. all code-combined ARQ transmissions of the same message. In this paper, we investigate a version we term *Fast-HARBINGER* where the sleep states are synchronized with the packet rate. Thus, the topology changes after each ARQ transmission. This provides an additional time-diversity benefit relative to Slow-HARBINGER at the expense of requiring more rapid sleep states.

II. THE HARBINGER PROTOCOL

We assume a two-dimensional ad hoc wireless network with Poisson distributed nodes. Due to a periodic sleep cycle, the density ρ of active nodes is much lower than the actual node density and the topology periodically changes at random. The *network coherence time* τ is the duration that the topology remains fixed. The first *network coherence interval* (NCI) is the range $t : \{0 \leq t < \tau\}$ while the i^{th} NCI is $t : \{(i-1)\tau \leq t < i\tau\}$. 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 but neglects the influence of interference and fading.

Each packet's header contains the location of the source and destination and is encoded with a low rate code for maximum error protection. Once a node has message to transmit, sends a request-to-send (RTS) signaling packet at the start of the next NCI in an effort to search for nearby nodes. 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 restart handshaking at the start of 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 one closest to the destination).

Once a relay has been selected, the source will transmit the data packet to it and the process will start over with the relay acting as the new source. The key distinction between HARBINGER and GeRaF is in what happens if there is not a relay within range. Under the GeRaF protocol, the system must start over again. All information from the initial transmission will be lost. With HARBINGER, hybrid-ARQ is used to combine information accumulated over multiple transmissions. Distant nodes outside of the source's first attempt transmission range R_1 but within the maximum ARQ transmission range R_M 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 the *Slow-HARBINGER* protocol, described more fully in [5], a slowly changing topology is assumed and thus several ARQ transmissions can be attempted before the topology changes. An additional time diversity benefit can be achieved by intentionally causing the topology to change after each ARQ transmission, which is the basis of *Fast-HARBINGER*.

In Fast-HARBINGER, the source encodes 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). The constant M is called the *rate constraint* of the hybrid-ARQ protocol [4]. Fast-HARBINGER transmits each message via a source communication session which spans at most M NCIs. During each NCI, the source broadcasts an RTS packet. The RTS packets are numbered 1 through M to indicate which data frame will be sent if a node replies with a CTS packet (the session *expires* after the M^{th} RTS packet is broadcast). All nodes within range $R_M > R_1$ will be able to decode the RTS packet and then must make a local decision to continue listening. This decision is based on its relative location and how many more data packets for the current message could be transmitted. If it is impossible for the node to decode the message before the session expires (i.e. it woke up too late), then it will go back to sleep. Otherwise, it will send a CTS packet and remain awake for the remainder of the session, even if it is not yet able to decode the message. Upon receiving the first CTS packet, the source will begin to transmit coded packets (one per NCI) until one of the relays is able to decode it and replies with an ACK.

Once a node decides to receive packets, it will keep every packet it receives 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. Once a specific relay is chosen, all the active nodes within the coverage area will flush their memory (discard previously received packets) and start a brand new session. In the following, we distinguish between two versions of Fast-HARBINGER, *A* and *B*. With Fast-HARBINGER-*A*, a new session begins immediately, regardless of whether

there is a geographically advantaged relay within range R_M . The problem with Fast-HARBINGER-A is that if there is no relay within range during the first NCI, then the maximum coverage range will be $R_{M-1} < R_M$. Fast-HARBINGER-B overcomes this problem by waiting to start the next session until a geographically-advantaged node awakens within range R_M , thereby taking full advantage of the dynamic range expansion feature of hybrid-ARQ.

III. A MATHEMATICAL FRAMEWORK

Due to the use of hybrid-ARQ, the *effective* coverage area of HARBINGER increases after each ARQ transmission. Let the effective coverage radius after the m^{th} transmission be R_m . Under the assumption of capacity-approaching channel coding, code-combining hybrid-ARQ, exponential path loss, and AWGN, this radius can be found by first finding the channel capacity after the m^{th} transmission [4]:

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

where $\frac{\mathcal{E}_s}{N_o}$ is the transmit signal to noise ratio, μ the path loss coefficient [6], K_0 the signal propagation coefficient [6], and d_m the propagation distance. 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

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

The radius R_m is the maximum transmission range, i.e. the value of d_m that satisfies (2) with equality. For analytical convenience, we assume $R_1 = 1$ and normalize $\{R_m\}$ with respect to R_1 , resulting in

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

As in [2], the source is located at coordinates $(D, 0)$ and the destination at $(0, 0)$. In GeRaF, the coverage area of the source is represented by a single circle of unit radius centered at $(D, 0)$. In contrast, the Hybrid-ARQ mechanism of HARBINGER allows the network to be covered by a set of M concentric coverage circles centered at $(D, 0)$. Every node inside circle O_m of radius R_m (with $R_1 = 1$) can be reached through at most m ARQ transmissions. For notational consistency, we also define circle O_0 which has radius $R_0 = 0$.

We further define $D\nu$ concentric circles centered at $(0, 0)$ with each circle Q_i having radius i/ν . The concentric circles $\{Q_i\}$ quantize the whole range of possible distances from the source or relays to the destination, e.g. $0 \sim D$, into $D\nu$ intervals with ν being the number of quantization interval per unit distance. The l^{th} quantization interval $\Delta_l = Q_l - Q_{l-1}$. Once a relay node is chosen to forward the message, it automatically becomes the source during

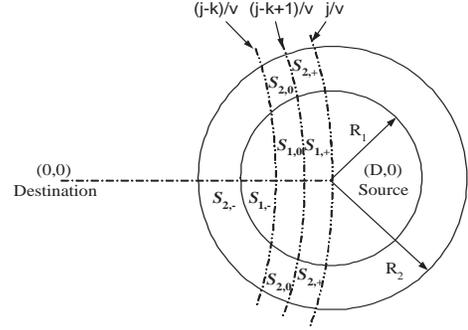


Fig. 1. Concentric coverage circles for HARBINGER with $M = 2$.

the next session of ARQ transmissions. Therefore, the distance between source and destination changes each time the message is decoded and the selected relay takes on the role of the new source.

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

$$\mathcal{S}_m = (O_m - O_{m-1}) \cap Q_j \quad \text{for } 1 \leq m \leq M \quad (4)$$

Any active node in \mathcal{S}_m could correctly decode the message by receiving exactly m data packets from the source. It is straightforward to show that

$$\begin{aligned} \cup_{m=1}^M \mathcal{S}_m &= O_M \cap Q_j \\ \mathcal{S}_i \cap \mathcal{S}_k &= \emptyset \quad \text{for } i \neq k \end{aligned} \quad (5)$$

Given $j - R_m\nu + 1 \leq l \leq j - R_{m-1}\nu$, \mathcal{S}_p is further divided into three disjoint regions by interval Δ_l for $p \geq m$. In particular,

$$\begin{aligned} \mathcal{S}_p &= \mathcal{S}_{p,+} \cup \mathcal{S}_{p,0} \cup \mathcal{S}_{p,-} \\ \mathcal{S}_{p,+} &= (O_p - O_{p-1}) \cap (Q_j - Q_l) \\ \mathcal{S}_{p,0} &= (O_p - O_{p-1}) \cap (Q_l - Q_{l-1}) \\ \mathcal{S}_{p,-} &= (O_p - O_{p-1}) \cap Q_{l-1} \end{aligned} \quad (6)$$

However, when $p < m$, \mathcal{S}_p is not further partitioned.

As a particular example, Fig. 1 shows the coverage partition in HARBINGER with rate constraint $M = 2$.

IV. PERFORMANCE ANALYSIS

When a node wakes up, it makes a localized decision to stay awake for the remainder of the session or to fall back asleep. Thus, the node distribution is time-dependent over the course of a particular session. Let X_i^t denote the event that partition \mathcal{S}_i contains at least one relay during the t^{th} NCI (a bar over an event denotes its complement). Likewise, $X_{i,\bullet}^t$ denotes the event that subpartition $\mathcal{S}_{i,\bullet}$ contains at least one potential relay during the t^{th} NCI, where \bullet could be '+', '0' or '-'.

A. Fast-HARBINGER-A

$\omega(j, k, b)$ denotes the joint probability of message progress from location $j/\nu, j > R_2\nu$ to interval Δ_{j-k+1} with a delay

of $b\tau, b \leq M$. In particular,

$$\omega(j, k, 1) = \begin{cases} Pr \{X_{1,0}^1 \cap \bar{X}_{1,-}^1\}, & k = 1, \dots, R_1\nu \\ 0, & k = R_1\nu + 1, \dots, R_2\nu \end{cases}$$

$$\omega(j, k, 2) = \begin{cases} Pr \{\bar{X}_1^1 \cap \bar{X}_{2,-}^1 \cap \bar{X}_{1,-}^2 \cap (X_{1,0}^2 \cup X_{2,0}^2)\}, & \text{for } k = 1, \dots, R_1\nu \\ Pr \{\bar{X}_1^1 \cap X_{2,0}^1 \cap \bar{X}_{2,-}^1\}, & \text{for } k = R_1\nu + 1, \dots, R_2\nu \end{cases}$$

where $\bar{X}_1^1 \cap \bar{X}_{2,-}^1 \cap \bar{X}_{1,-}^2 \cap (X_{1,0}^2 \cup X_{2,0}^2)$ denotes the event that the message progresses to $\mathcal{S}_{1,0} \cup \mathcal{S}_{2,0}$ with delay 2τ . In addition, $\omega_0(j) = Pr \{\bar{X}_1^1 \cap \bar{X}_2^1 \cap \bar{X}_1^2\}$ denotes the probability that no progress is made within the current session, and thus the same message has to be retransmitted in the next session.

More generally, as $M > 2, j > R_M\nu$ (i.e. the destination is outside coverage area O_M), and $k = R_{p-1}\nu + 1, \dots, R_p\nu$, the corresponding joint probability of message progress $\omega(j, k, b)$ is found as

$$\omega(j, k, b) = \begin{cases} Pr \{\Omega(b, \cup_{i=p}^b \mathcal{S}_{i,0})\} & \text{for } b = p, \dots, M \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where

$$\Omega(b, \cup_{i=p}^b \mathcal{S}_{i,0}) = (\cap_{i=1}^{b-1} \cap_{t=1}^{b-i} \bar{X}_i^t) \cap (\cap_{i=p}^b \bar{X}_{i,-}^{b+1-i}) \cap (\cup_{i=p}^b X_{i,0}^{b+1-i}) \quad (8)$$

$\Omega(b, A)$ denotes the event that the message advances to region A with delay $b\tau$. The probability of these independent Poisson random events can be found as

$$Pr \{\cap_{i=1}^{b-1} \cap_{t=1}^{b-i} \bar{X}_i^t\} = \prod_{i=1}^{b-1} \exp\{-\rho A(j/\nu, j/\nu, R_i)\}$$

$$Pr \{(\cap_{i=p}^b \bar{X}_{i,-}^{b+1-i}) \cap (\cup_{i=p}^b X_{i,0}^{b+1-i})\} = \exp\{-\rho A(j/\nu, (j-k)/\nu, R_b)\} - \exp\{-\rho A(j/\nu, (j-k+1)/\nu, R_b)\} \quad (9)$$

In addition,

$$\omega_0(j) = Pr \{\cap_{i=1}^M \cap_{t=1}^{M+1-i} \bar{X}_i^t\}$$

$$= \prod_{i=1}^M \exp\{-\rho A(j/\nu, j/\nu, R_i)\} \quad (10)$$

On the other hand, when $j = R_{m-1}\nu + 1, \dots, R_m\nu$, the destination is within partition \mathcal{S}_m , thus is reachable with delay at most $m\tau$. Since in this case the destination is always reachable, $w_0(j) = 0$. When $k = R_{p-1}\nu + 1, \dots, R_p\nu$, and $p < m$

$$\omega(j, k, b) = \begin{cases} Pr \{\Omega(b, \cup_{i=p}^b \mathcal{S}_{i,0})\}, & \text{for } b = p, \dots, m-1 \\ 0, & \text{otherwise} \end{cases}$$

When $k = R_{m-1}\nu + 1, \dots, j$,

$$\omega(j, k, b) = Pr \{\cap_{l=1}^{b-1} \cap_{t=1}^{b-l} \bar{X}_l^t\} \delta(j, k) \delta(m, b) \quad (11)$$

Correspondingly, by using the same recursive approach in [2] the upper and lower bounds of average delay can be computed as

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

$$n_2(j) = \sum_{k=1}^{R_M\nu} \sum_{b=1}^M \omega(j, k, b) (n_2(j-k) + b) + \omega_0(j) (n_2(j) + M) \quad (12)$$

with initial condition $n_1(j) = n_2(j) = 1$ when $j = 1, \dots, \nu$ and $n_1(j) = n_2(j) = 2$ when $j = \nu + 1, \dots, R_2\nu$.

Under the assumption that most of the energy is sent transmitting data (e.g. the energy used to transmit control messages is negligible), then energy efficiency can be quantified in terms of the average number of ARQ data packet transmissions per end-to-end message. Now consider a joint probability $\omega(j, k, b, l)$ denoting the probability of message progress from location j/ν to interval Δ_{j-k+1} with delay $b\tau$ and l data packet transmission ($l \leq b$). Then,

$$\omega(j, k, b) = \sum_{l=1}^b \omega(j, k, b, l) \quad (13)$$

$\omega(j, k, b, l)$ can be found as

$$\omega(j, k, 1, 1) = \omega(j, k, 1), \quad \text{for } k = 1, \dots, R_2\nu$$

$$\omega(j, k, 2, 1) = \begin{cases} Pr \{\varphi_2\}, & k = 1, \dots, R_1\nu \\ 0, & k = R_1\nu + 1, \dots, R_2\nu \end{cases}$$

$$\omega(j, k, 2, 2) = \begin{cases} Pr \{\varphi_1 - \varphi_2\}, & \text{for } k = 1, \dots, R_1\nu \\ Pr \{\bar{X}_1^1 \cap X_{2,0}^1 \cap \bar{X}_{2,-}^1\}, & \text{for } k = R_1\nu + 1, \dots, R_2\nu \end{cases} \quad (14)$$

where

$$\varphi_1 = \bar{X}_1^1 \cap \bar{X}_{2,-}^1 \cap \bar{X}_{1,-}^2 \cap (X_{1,0}^2 \cup X_{2,0}^2)$$

$$\varphi_2 = \bar{X}_1^1 \cap \bar{X}_2^1 \cap X_{1,0}^2 \cap \bar{X}_{1,-}^2 \quad (15)$$

More generally, as $M > 2, \omega(j, k, b, l)$ could be computed as follows. When integer $j > R_M\nu$, and $k = R_{p-1}\nu + 1, \dots, R_p\nu$, the joint probability $\omega(j, k, b, l)$ is found as

$$\omega(j, k, b, l) = \begin{cases} Pr \{b, l, \Omega(\cup_{i=p}^b \mathcal{S}_{i,0}) - \Omega(b, l-1, \cup_{i=p}^b \mathcal{S}_{i,0})\}, & \text{for } b = p, \dots, M \text{ and } l = p, \dots, b \\ 0, & \text{otherwise} \end{cases}$$

where

$$\Omega(b, l, \cup_{i=p}^b \mathcal{S}_{i,0}) = (\cap_{t=1}^{b-l} \cap_{i=1}^M \bar{X}_i^t) \cap (\cap_{t=b+1-l}^{b-1} \cap_{i=1}^{b-t} \bar{X}_i^t) \cap (\cap_{i=p}^l \bar{X}_{i,-}^{b+1-i}) \cap (\cup_{i=p}^l X_{i,0}^{b+1-i}) \quad (16)$$

$\Omega(b, l, A)$ denotes the joint event that the message advance to the region A with exactly $b\tau$ delay and at most l data packets in the session. It is straightforward that

$$\Omega(b, b, \cup_{i=p}^b \mathcal{S}_{i,0}) = \Omega(b, \cup_{i=p}^b \mathcal{S}_{i,0}) \quad (17)$$

The probability of these independent Poisson random events can be found as

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

$$Pr \left\{ \bigcap_{t=b+1-l}^{b-1} \bigcap_{i=1}^{b-t} \bar{X}_i^t \right\} = \prod_{i=1}^{l-1} \exp \{-\rho A(j/\nu, j/\nu, R_i)\} \quad (18)$$

and $Pr \left\{ \left(\bigcap_{i=p}^l \bar{X}_{i,-}^{b+1-i} \right) \cap \left(\bigcup_{i=p}^l X_{i,0}^{b+1-i} \right) \right\}$ is defined in (9).

On the other hand, when $j = R_{m-1}\nu + 1, \dots, R_m\nu$, the destination is within partition \mathcal{S}_m , thus the source will always transmit data packet in each NCI. Therefore, $\omega(j, k, b, l)$ becomes

$$\omega(j, k, b, l) = \omega(j, k, b)\delta(b, l) \quad (19)$$

Correspondingly, the upper and lower bounds of average number of data packets per message can be calculated recursively using

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

$$e_2(j) = \sum_{k=1}^{R_M\nu} \sum_{b=1}^M \sum_{l=1}^b \omega(j, k, b, l)(e_2(j-k) + l) + \omega_0(j)e_2(j) \quad (20)$$

with initial condition $e_1(j) = e_2(j) = 1$ when $j = 1, \dots, \nu$ and $e_1(j) = e_2(j) = 2$ when $j = \nu + 1, \dots, R_2\nu$.

B. Fast-HARBINGER-B

With Fast-HARBINGER-B, each session begins only when there is a geographically advantaged node within range R_M . In this case, the joint probability of message delay is equivalent to the the joint probability of ARQ re-transmissions, i.e. $\omega(j, k, b) = \omega(j, k, b, b)$. More specifically, when $j > R_M\nu$, and $k = R_{p-1}\nu + 1, \dots, R_p\nu$,

$$\omega(j, k, b, l) = \omega(j, k, b, b)\delta(b, l)$$

$$\omega_0(j) = Pr \left\{ \bigcap_{i=1}^M \bar{X}_i^1 \right\}$$

$$= \exp \{-\rho A(j/\nu, j/\nu, R_M)\} \quad (21)$$

On the other hand, when $R_{m-1}\nu + 1 \leq j \leq R_m\nu$, the destination is within partition \mathcal{S}_m and is reachable with at most $m\tau$ delay. Therefore $w_0(j) = 0$ and $\omega(j, k, b, l)$ could be computed by (19). The upper and lower bounds of average delay can be computed as

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

$$n_2(j) = \sum_{k=1}^{R_M\nu} \sum_{b=1}^M \omega(j, k, b, b)(n_2(j-k) + b) + \omega_0(j)(n_2(j) + 1) \quad (22)$$

with initial condition $n_1(j) = n_2(j) = 1$ when $j = 1, \dots, \nu$ and $n_1(j) = n_2(j) = 2$ when $j = \nu + 1, \dots, R_2\nu$.

Likewise, the upper and lower bounds of average number of data packets per message becomes

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

$$e_2(j) = \sum_{k=1}^{R_M\nu} \sum_{b=1}^M \omega(j, k, b, b)(e_2(j-k) + b) + \omega_0(j)e_2(j) \quad (23)$$

with initial condition $e_1(j) = e_2(j) = 1$ when $j = 1, \dots, \nu$ and $e_1(j) = e_2(j) = 2$ when $j = \nu + 1, \dots, R_2\nu$.

V. NUMERICAL RESULTS

The analytical bounds on performance are plotted in Fig.2-5 for different values of rate-constraint M and source/destination separation D . The message delay is normalized by network (topology) coherence time τ . The per-packet code rate is $r = 1$ and the number of intervals per unit distance $\nu = 50$. As observed in Fig.2, the upper and lower bounds are quite close to each other indicating the tightness of both bounds. In fact, as the interval spacing $1/\nu$ becomes smaller, the bounds will become more accurate and vice versa. When $M = 1$, HARBINGER is equivalent to GeRaF. In a relatively dense network, HARBINGER has the same performance as GeRaF. In fact, both GeRaF and HARBINGER will asymptotically converge to a message delay of $\lfloor D + 1 \rfloor$ as node density $\rho \rightarrow \infty$. The real benefit of Fast-HARBINGER is in sparse networks where $\rho \rightarrow 0$, indicating that HARBINGER will allow nodes to remain in a sleep state for a relatively longer percentage of time than GeRaF (for the same total node density). As observed in Fig.2, in low density networks, Fast-HARBINGER significantly reduces the message delay as M increases. In fact, slightly increasing the rate constraint from $M = 1$ to $M = 2$ reduces the delay by almost 50%. However, as we further increase M , diminishing returns kicks in very fast.

Fig. 3 indicates the average number of data packet transmissions per message. As we have already demonstrated that both bounds are tight, we show only the lower bound in Fig.3. Assuming that data packets consume much more energy than signaling packets, then Fig.3 demonstrates the energy efficiency of both GeRaF and Fast-HARBINGER protocols. As observed in Fig.3, Fast-HARBINGER actually consumes more energy than GeRaF primarily due to its relatively aggressive packet transmission strategy and non-linear coverage expansion behavior. As $\rho \rightarrow \infty$, both GeRaF and Fast-HARBINGER asymptotically require $\lfloor D + 1 \rfloor$ data packet transmissions for each message. Further notice that unlike the message delay, which decreases as the rate constraint increases, the energy consumption of Fast-HARBINGER actually increases with a higher rate constraint. As the rate constraint gets fairly large, i.e. $M = 12$, the message delay is almost equiv-

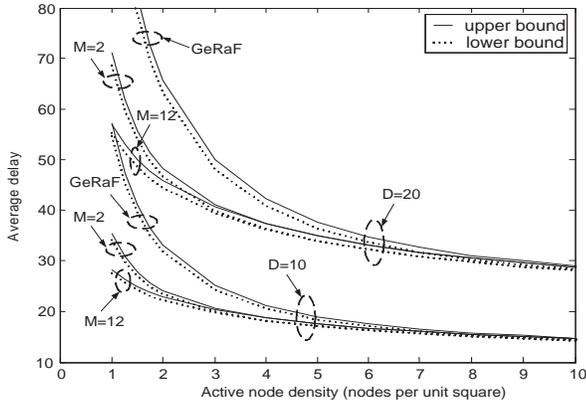


Fig. 2. The average delay of Fast-HARBINGER-A under different rate constraints M . $M = 1$ corresponds to GeRaF.

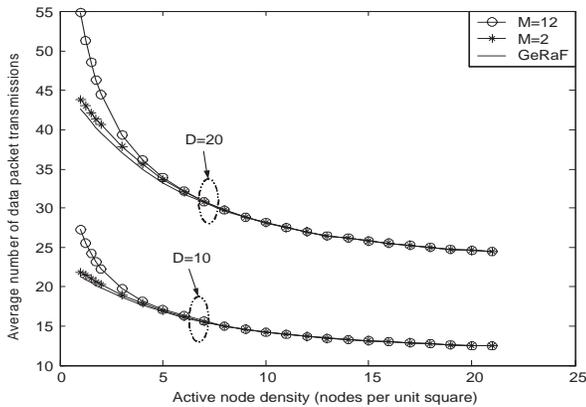


Fig. 3. Lower bound on average data packet retransmissions per message in Fast-HARBINGER-A under different rate constraints M .

alent to the average data packet transmission, indicating that there always exists at least one relay node within the coverage area at the first NCI of each session. In practice, energy consumption must be traded off for delay and so $M = 2$ or 3 might be an appropriate number for network implementation. Notice in Fig.4 and Fig.5 that although Fast-HARBINGER-B has almost the same energy efficiency as that of Fast-HARBINGER-A, it manages to further reduce the average delay by about 10%, primarily due to its ability to fully exploit the coverage expansion mechanism. Therefore, Fast-HARBINGER-B is more desirable in practice. Finally, we observe that both message delay and energy efficiency of the HARBINGER protocol is linearly proportional to the separation distance as indicated in Fig. 2 and Fig. 3.

VI. 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 low rate constraint, indicating that a smaller duty-cycle sleep schedule could be used with HARBINGER, thereby increasing the useful lifetime of sensor networks.

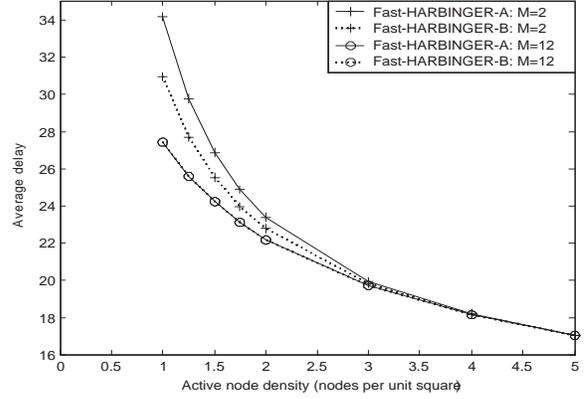


Fig. 4. Delay comparison of Fast-HARBINGER-A and Fast-HARBINGER-B under different rate constraints M , and $D = 10$.

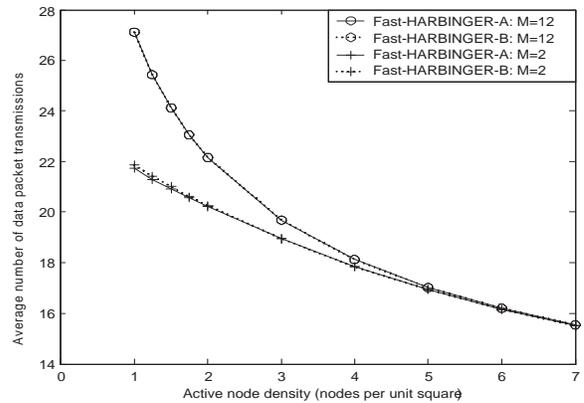


Fig. 5. The energy efficiency of Fast-HARBINGER-A and Fast-HARBINGER-B under different rate constraints M , and $D = 10$.

Alternatively, for the same sleep schedule, HARBINGER allows reduced end-to-end delay compared to GeRaF. In contrast with Slow-HARBINGER [5], Fast-HARBINGER intentionally changes the network topology prior to each data packet transmission, thereby achieving an additional time diversity benefit in time varying networks. Such benefit is primarily demonstrated by its significant reduction in the message delay.

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] M. C. Valenti and B. Zhao, "Hybrid ARQ-based intra-cluster geographically-informed relaying," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, (Monterey, CA), Nov. 2004. to appear.
- [6] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice Hall PTR, second ed., 2002.