

# On energy efficiency of geographic opportunistic routing in lossy multihop wireless networks

Kai Zeng · Jie Yang · Wenjing Lou

© Springer Science+Business Media, LLC 2012

**Abstract** Geographic opportunistic routing (GOR) is an emerging technique that can improve energy efficiency in lossy multihop wireless networks. GOR makes local routing decision by using nodes' location information, and exploits the broadcast nature and spatial diversity of the wireless medium to improve the packet forwarding reliability. In this paper, our goal is to fully understand the principles and tradeoffs in GOR, thus provide insightful analysis and guidance to the design of more efficient routing protocols in multihop wireless networks. We propose a local metric, one-hop energy efficiency (OEE), to balance the packet advancement, reliability and energy consumption in GOR. We identify and prove important properties about GOR on selecting and prioritizing the forwarding candidates in order to maximize the expected packet advancement. Leveraging the proved properties, we then propose two localized candidate selection algorithms with  $O(N^3)$  running time to determine the forwarding candidate set that maximizes OEE, where  $N$  is the number of available next-hop neighbors. Through extensive simulations, we show that GOR applying OEE achieves better energy efficiency than the existing geographic routing and

blind opportunistic routing schemes under different node densities and packet sizes.

**Keywords** Multihop wireless networks · Geographic opportunistic routing · Energy efficiency

## 1 Introduction

Many recent experimental studies on multihop wireless networks (including wireless ad hoc and sensor networks) have shown that wireless links can be highly unreliable [1–3]. Recently, a new routing paradigm, known as opportunistic routing (OR) [4–6] (or contention-based routing [7]), was proposed to cope with the unreliability of link quality in such networks.

Different from traditional routing (TR), where a single next-hop node is predefined for each packet forwarding at the network layer, OR selects multiple neighboring nodes as forwarding candidates at network layer and chooses one candidate that successfully received the packet as the actual forwarder at the MAC layer. In other words, TR makes routing decision *before* each actual physical layer transmission, while OR does it *after* each physical transmission according to the instant packet reception results at the neighboring nodes.

The philosophy behind OR is to improve the transmission reliability by utilizing the spatial diversity of the wireless medium to obtain better performance (such as energy efficiency, reliability and throughput). The intuition is that owing to the broadcast nature and spatial diversity of the wireless medium, the probability of at least one forwarding candidate correctly receiving the packet will increase when more forwarding candidates are involved. Thus, the retransmission overhead is reduced and the

---

K. Zeng  
Department of Computer and Information Science,  
University of Michigan, Dearborn, MI 48128, USA  
e-mail: kzeng@umich.edu

J. Yang  
Department of Computer Science, University of California,  
Davis, CA 95616, USA  
e-mail: jieya@ucdavis.edu

W. Lou (✉)  
Department of Computer Science, Virginia Polytechnic Institute  
and State University, Blacksburg, VA 24061, USA  
e-mail: wjlou@vt.edu

packet will get more chance to be forwarded towards the destination for each transmission. The reduced retransmission not only reduces the transmission and reception energy consumption on nodes, but also alleviates the medium contention. Therefore, OR can improve energy efficiency and throughput.

One type of OR schemes, geographic opportunistic routing (GOR) [4, 5, 7, 8], uses nodes' location information to define the forwarding candidate set and prioritize candidates. The advantage of GOR over link state based OR (such as ExOR [1] and least-cost opportunistic routing (LCOR) [9]) lies in several aspects. For one, nodes need to know only the local information (location and/or link quality) of their one-hop neighbors in order to forward packets and hence the state stored is minimum. Further, GOR conserves energy and bandwidth since discovery floods and state propagation are not required beyond a single hop.

In this paper, we focus on the energy efficiency issue about GOR in wireless sensor networks (WSNs). In WSNs, the major portion of energy consumption comes from packet transmission and reception [10]. Although GOR can improve energy efficiency by reducing the MAC layer retransmission, it involves more forwarding candidates than traditional geographic routing (GR). Therefore, it consumes more reception energy. There exist tradeoffs between packet advancement, reliability and energy consumption in GOR. Unfortunately, the analysis of this tradeoff is missing in the literature.

In this paper, we propose a local metric, One-hop Energy Efficiency (OEE), to strike a good balance between the benefit (packet advancement and reliability) and the cost (energy consumption) in GOR. We identify and prove important properties about GOR on selecting and prioritizing the forwarding candidates in order to maximize the expected packet advancement (EPA). Leveraging the proved properties, we then propose two localized candidate selection algorithms to determine the forwarding candidate set that maximizes OEE. These two algorithms both achieve  $O(N^3)$  running time, where  $N$  is the number of available next-hop neighbors. Through extensive simulations, we show that GOR applying OEE achieves better energy efficiency than the existing geographic routing and opportunistic routing schemes. We would like to emphasize that although we mainly focus on the energy efficiency issue, which is a major concern in sensor networks, the framework and analysis presented in this paper are expected to be extensible and applicable to other multihop wireless networks and be useful when analyzing other performance such as throughput.

The rest of the paper is organized as follows. The related work is introduced in Sect. 2. Section 3 proposes the definitions of EPA and OEE, and formulates the energy-

efficient GOR problem. Several important properties about the EPA and the corresponding forwarding candidate set are revealed and analyzed in Sect. 5. Two efficient localized candidate selection algorithms are proposed in Sect. 6. Simulation results are presented in Sect. 7. Conclusions are drawn in Sect. 8.

## 2 Related work

Our work mainly relates to the following two research topics: geographic routing and opportunistic routing.

### 2.1 Geographic routing

In geographic routing [11–19], the forwarding node is aware of the location information of its own, its neighbors, and the destination, thus a packet can be effectively forwarded towards the direction of its destination. The node location information can be obtained by prior configuration, by the Global Positioning System (GPS) receiver, or through some sensor self-configuring localization mechanisms [20, 21]. As each node only needs to maintain its neighbors' information, the properties such as good scalability, statelessness, and low maintenance overhead make geographic routing an efficient technique especially in large-scale wireless sensor networks.

Several recent experimental studies on wireless ad-hoc and sensor networks [1, 2] have shown that wireless links can be highly unreliable and this factor must be explicitly taken into account when devising higher-layer protocols. More recent works on geographic routing therefore are more focused on this practical lossy channel situation. Seada, et al. [17] articulated the distance–hop energy tradeoff for geographic routing. They concluded that packet advancement timing packet reception ratio, is an optimal metric for making localized geographic routing decisions in lossy wireless networks with ARQ (Automatic Repeat reQuest) mechanisms, and is also a good metric for No-ARQ scenarios. Zorzi and Armaroli also independently proposed the same link metric [22]. Lee et al. [18] presented a more general framework called normalized advance (NADV) to normalize various types of link cost such as transmission times, delay and power consumption. Unfortunately, these metrics only apply to geographic routing which involves single forwarding candidate and cannot be directly applied to GOR.

### 2.2 Opportunistic routing

Opportunistic routing integrates the network and MAC layers. At the network layer, a set of forwarding candidates or a region are selected. At the MAC layer, one node is

chosen as the forwarder depending on the node’s availability and wireless link quality. Several MAC rendezvous [23, 24] and coordination schemes [4, 7, 25] have been proposed to coordinate nodes involved in the OR and ensure a single relay is chosen to forward the packet at each hop.

There are generally two types of OR, location-aided [4, 5, 7, 8] and link state based ones [6, 9]. In this paper, we focus on the former, the geographic opportunistic routing (GOR).

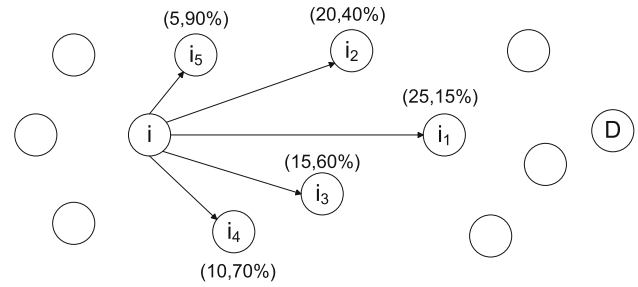
One of the most detailed works in GOR is GeRaF [4] where candidate forwarders that are geographically closer to the destination than the current forwarding node are divided into sets of priority regions with nodes closer to the destination having higher relay priorities. One similar work to [4] is [5] where the network layer specifies a set of nodes by defining a forwarding region in space that consists of the candidate nodes and the data link layer selects the first node available from that set to be the next hop node. A delay constraint hop counts optimization scheme with adaptive node wake-up discipline is studied in this work. [7] proposed contention-based forwarding for mobile ad hoc networks and discussed three suppression strategies to avoid packet duplication. [26] compares the performances such as power consumption, delay and packet delivery ratio of greedy geographic routing with region based OR. [27] proposes a framework to model OR in low traffic scenarios.

All of these GOR schemes blindly involve all the available next-hop neighbors as forwarding candidates. None of these GOR achieves good balance between packet advancement, reliability and energy consumption. None of them provide a thorough understanding of how the selection of the forwarding candidate set will affect the energy efficiency. Although a recent work on GOR studied the throughput and delay performance of GOR by selecting a subset of available next-hop neighbors as forwarding candidates, energy efficiency is not its concern [8]. The questions, such as “(a) how many and which neighbor nodes should be involved in the local opportunistic forwarding?”; and “(b) what criteria should affect the relay priority among the forwarding candidates?”, remain unanswered. In this paper, we will answer these questions with energy efficiency as a major concern.

### 3 System and energy consumption model

#### 3.1 System model

In this paper, we assume the local GOR scenario as in Fig. 1. Assuming node  $i$ , i.e., the transmitter, is forwarding a packet to a sink/destination  $D$ .  $i_j$  is one of  $i$ ’s neighbors which are closer to  $D$  than  $i$ . Let  $\mathcal{C}$  be the set of  $i_j$ . We call  $\mathcal{C}$



**Fig. 1** Example in which node  $i$  is forwarding a packet to a remote destination  $D$

as the *available next-hop node set* of node  $i$ . Let  $N = |\mathcal{C}|$  denote the number of nodes in  $\mathcal{C}$ .

Each  $i_j$  corresponds to a pair,  $(d_j, p_j)$ , where  $d_j$  defined in Eq. (1) is the packet advancement to the destination when packet sent by  $i$  is relayed by  $i_j$ , and  $p_j$  is the packet reception ratio (PRR) from node  $i$  to  $i_j$ . Note that  $d_j$  is the advancement, not the distance between the node  $i$  and its neighbor  $i_j$ .

$$d_j = \text{Dist}(i, D) - \text{Dist}(i_j, D) \tag{1}$$

where  $\text{Dist}(i, D)$  and  $\text{Dist}(i_j, D)$  are the Euclidean distance between  $i$  and  $D$  and between  $i_j$  and  $D$ , respectively.

A node is a neighbor of  $i$  when PRR from  $i$  to it is larger than some non-negligible probability threshold  $p_T$ <sup>1</sup>. From the definition of  $d_j$  and  $p_j$ , we know that  $d_j > 0$  and  $p_T < p_j < 1$ . The PRR information on each link can be obtained by using probe messages [28] or estimated by Signal-to-Noise Ratio (SNR) [3]. We assume the PRR on each link is independent. Let  $\mathcal{F}$  denote the forwarding candidate node set of node  $i$ , which includes all the nodes selected to get involved in the local opportunistic forwarding. Let  $r = |\mathcal{F}|$  denote the number of nodes in  $\mathcal{F}$ . Here  $\mathcal{F}$  is a subset of  $\mathcal{C}$ , while in the existing OR schemes [4, 6],  $\mathcal{F} = \mathcal{C}$ . We assume all the nodes in  $\mathcal{C}$  and  $\mathcal{F}$  are descending ordered according to the advancement s.t.  $d_m \geq d_n, \forall m < n$ .

The GOR procedure is described as follows. A sender, node  $i$ , selects  $\mathcal{F}$  based on its knowledge of  $\mathcal{C}$  ( $d_j$ ’s and  $p_j$ ’s) and some performance goal (e.g. energy efficiency). When the forwarding candidates in  $\mathcal{F}$  receive the packet broadcasted by  $i$ , they follow a specific priority to relay the packet. That is, the first-priority candidate should attempt to relay the packet when it received it correctly. If it failed to receive the packet, the second-priority candidate will claim the forwarding responsibility if it received the packet successfully, otherwise the third-priority one, etc. In short, a forwarding candidate will only relay the packet if all the nodes with higher priorities failed to do so. The actual forwarder will become a new transmitter and suppress all

<sup>1</sup> In this paper,  $p_T = 0.1$ .

the other potential forwarders in  $\mathcal{F}$ . When no forwarding candidate has successfully received the packet, the sender will retransmit the packet if retransmission is enabled. The sender will drop the packet when the retransmissions reach the limit. This procedure iterates until the packet arrives at the destination. When the packet gets stuck due to a communication void where no available next-hop neighbors are closer to the destination than the current node, the packet will be dropped. Mechanisms such as FACE routing [14] or perimeter forwarding in GPSR [15] can be applied here to deal with the communication void problem, but it is beyond the scope of this paper.

In this paper, we focus ourselves on greedy forwarding, and aim to achieve energy efficiency by locally selecting and prioritizing forwarding candidates in a judicious way.

### 3.2 Energy consumption model

In wireless networks, the energy consumption from communication mainly comes from three parts: transmission, reception, and idle listening [29]. For GOR, packet forwarding is a random event due to its opportunistic nature. So we use

$$\bar{E}_{or} = \bar{E}_i + \bar{E}_{\mathcal{F}} + \bar{E}_o \quad (2)$$

to represent the expected energy consumption for one local opportunistic forwarding.  $\bar{E}_i$ ,  $\bar{E}_{\mathcal{F}}$ , and  $\bar{E}_o$  are the expected energy consumption on transmitter  $i$ , all nodes in the candidate set  $\mathcal{F}$ , and other neighboring nodes of node  $i$ , respectively. Each of these three energy consumption consists of transmission, reception, and idle listening parts. Note that  $\bar{E}_i$ ,  $\bar{E}_{\mathcal{F}}$ , and  $\bar{E}_o$  depend on many factors, such as neighborhood size, forwarding candidate set, forwarding priority, MAC protocol, link quality, node wakeup/sleep pattern, packet size, data rate, transmission power, receiving power, and idle listening power, etc. The energy consumption for delivering one packet from a source  $S$  to a destination  $D$  is the total energy consumption of all the nodes involved into the OR.

## 4 Local metric: opportunistic energy efficiency (OEE)

Recall that our goal is to approach a minimum energy consumption to delivery the packet from the source to destination by locally selecting forwarding candidate set at each hop. In this section, we will first introduce the metric of EPA to describe how much closer a packet can be forwarded towards the destination per transmission in GOR given a forwarding candidate set  $\mathcal{F}$ . We then examine the expected energy consumption of one-hop local opportunistic forwarding given relatively ideal MAC and wakeup/sleep mechanism. We finally propose our a new local

metric: opportunistic energy efficiency OEE), which indicates one-hop expected packet advancement per unit of energy consumption.

### 4.1 Expected packet advancement (EPA)

Recall that our goal is to achieve the tradeoff between the packet advancement, reliability and energy consumption. We now introduce the metric of EPA to describe how much closer a packet can be forwarded towards the destination per transmission in GOR.

Let  $\pi_j(\mathcal{F}) = \langle i_{j_1}, i_{j_2}, \dots, i_{j_r} \rangle$  be one permutation of nodes in  $\mathcal{F}$ , and the order indicates that nodes will attempt to forward the packet with priority  $i_{j_1} > i_{j_2} > \dots > i_{j_r}$ . We define the EPA for the ordered forwarding candidate set  $\pi_j(\mathcal{F})$  in one opportunistic forwarding attempt in GOR as in Eq. (3).

$$\text{EPA}(\pi_j(\mathcal{F})) = \sum_{k=1}^r d_{j_k} p_{j_k} \cdot \prod_{n=0}^{k-1} (1 - p_{j_n}) \quad (3)$$

where  $p_{j_0} := 0$ . The physical meaning of Eq. (3) is the expected packet advancement achieved by GOR using the ordered forwarding candidate set  $\pi_j(\mathcal{F})$ . When  $r = 1$ , Eq. (3) degenerates to the metric of *advancement*  $\times$  *PRR* in traditional geographic routing [17, 18].

### 4.2 Opportunistic forwarding MAC

Note that  $\bar{E}_{or}$  depends on many factors, such as MAC and wakeup/sleep patterns. For different opportunistic forwarding coordination protocols and sleep/wakeup patterns, the energy consumption in the network should be different. In this paper, in order to get more insight about the impact of candidate selection on the energy efficiency of GOR, we assume the following relatively ideal coordination scheme and wake-up mechanism.

Assuming all the nodes of node  $i$  is sleeping in the beginning of a local opportunistic forwarding. When node  $i$  wants to forward a data packet to its forwarding candidates, it uses a low-power wakeup radio to wakeup the desired forwarding candidate(s) [30]. In the wakeup packet, the address or ID of the forwarding candidate(s) are included. The sequence of the address represents the forwarding priority. When receiving the wakeup packet, each neighboring node will exam if it is in the forwarding candidate set. If not, it will keep sleeping. Otherwise, it will wake up in order to receive the subsequent data packet. Assuming all the forwarding candidate nodes can be woken up reliably and fast slotted acknowledgement (FSA) [25] is used to coordinate the opportunistic forwarding. After sensing the channel has been idle for a DIFS (distributed inter-frame space), the transmitter locally broadcasts the data

packet. Each candidate waits for a short time before deciding whether it should broadcast ACK. Candidates with lower priority will wait longer. If the first-priority candidate receives the packet correctly, with a time delay of SIFS (short inter-frame space), it will send out an ACK. The ACK is used for acknowledging the sender of the reception as well as suppressing other candidates from duplicate forwarding. Meanwhile, all the other candidates sense the channel for a short period of time to tell whether they detect this ACK. If the answer is positive, they stop sensing the channel and simultaneously suppress their own attempts of sending ACK and forwarding the data packet. Otherwise, they assume the first-priority candidate missed the data packet and the second-priority candidate takes the responsibility of sending ACK and all the remaining lower priority candidates continue to monitor this second ACK. The coordination process goes on until some successful candidate finally sends out an ACK. By receiving the ACK, the transmitter and forwarding candidates will go back to sleep. If none of the candidates successfully received the packet, the sender will retransmit that packet after a timeout (e.g., 3SIFS + ACK time). It has shown in [25] that FSA is very efficient and can ensure the forwarding priority and avoid duplicate forwarding almost for sure.

### 4.3 Expected energy consumption of one-hop opportunistic forwarding

Under the above wakeup mechanism and coordination protocol, we can decide the energy consumption at transmitter and candidate sides.

The power consumption of the low-power wakeup circuit is about  $2.6 \mu W$  [30], which is negligible comparing to the power consumption of data transmission/reception and idle listening. That is, we assume  $\bar{E}_o = 0$ .

At transmitter side, we have the following expected energy consumption:

$$\bar{E}_i = \bar{E}_i^{tx} + \bar{E}_i^{rx} + \bar{E}_i^{idle} \tag{4}$$

where

$$\bar{E}_i^{tx} = \frac{1}{P_{\mathcal{F}}} E_{data\_tx} \tag{5}$$

$$\bar{E}_i^{rx} = E_{ACK\_rx} \tag{6}$$

$$\begin{aligned} \bar{E}_i^{idle} \approx & \frac{1}{P_{\mathcal{F}}} (E_{to} + p_1 E_{SIFS} + 2p_2(1 - p_1) E_{SIFS} \\ & + \dots + rp_r \prod_{j=1}^{r-1} (1 - p_j) E_{SIFS}) \end{aligned} \tag{7}$$

where  $P_{\mathcal{F}}$  is the probability of at least one forwarding candidate correctly receiving the packet. So  $\frac{1}{P_{\mathcal{F}}}$  is the number of expected transmissions the transmitter will make and Eq. (5) represents the expected energy

consumption of transmission at transmitter. Equation (6) says for a successful data forwarding, there is one ACK received by the transmitter. For representation simplicity, in Eq. (7), we assume the transmitter always wait  $t_o$  time and cost  $E_{to}$  energy before transmission. Then for each transmission (first time or retransmission) from the transmitter, the transmitter has to wait SIFS time and cost  $E_{SIFS}$  with probability  $p_1$  if node  $i_1$  correctly receives the packet. It then has to wait 2SIFS and cost  $2E_{SIFS}$  with probability  $p_2(1 - p_1)$  if node  $i_1$  does not correctly receive the packet but  $i_2$  does, etc.

At the forwarding candidate side, we have the following total expected energy consumption in the set  $\mathcal{F}$ :

$$\bar{E}_{\mathcal{F}} = \bar{E}_{\mathcal{F}}^{tx} + \bar{E}_{\mathcal{F}}^{rx} + \bar{E}_{\mathcal{F}}^{idle} \tag{8}$$

where

$$\bar{E}_{\mathcal{F}}^{tx} = E_{ACK\_tx} \tag{9}$$

$$\bar{E}_{\mathcal{F}}^{rx} = (r - 1)E_{ACK\_rx} + \frac{r}{P_{\mathcal{F}}} E_{data\_rx} \tag{10}$$

$$\begin{aligned} \bar{E}_{\mathcal{F}}^{idle} \approx & \frac{r}{P_{\mathcal{F}}} (E_{to} + p_1 E_{SIFS} + 2(1 - p_1)p_2 E_{SIFS} \\ & + \dots + rp_r \prod_{j=1}^{r-1} (1 - p_j) E_{SIFS}) \end{aligned} \tag{11}$$

Equation (9) says one of the forwarding candidates will eventually send an ACK when the forwarding is successful. Equation (10) indicates that when one forwarding candidate transmits the ACK, the other  $r - 1$  candidates receive it. It also indicates that for each data transmission from the transmitter, there are  $r$  receptions at the forwarding candidates. Equation (11) is similar to Eq. (7), except that we time  $r$  in the front. Because the idle listening time are the same at transmitter and candidates side, but we have  $r$  candidates.

By substituting Eqs. (5, 6), and (7) into Eq. (4), Eqs. (9, 10), and (11) into Eq. (8), and then Eq. (7) and Eq. (8) into Eq. (2), with  $\bar{E}_o = 0$ , Eq. (2) can be represented as

$$\begin{aligned} \bar{E}_{or} = & \frac{1}{P_{\mathcal{F}}} [E_{data\_tx} + rE_{data\_rx} \\ & + (r + 1)(E_{to} + \sum_{i=1}^r i \cdot E_{SIFS} \cdot p_i \prod_{j=1}^{i-1} (1 - p_j))] \\ & + E_{ACK\_tx} + rE_{ACK\_rx} \end{aligned} \tag{12}$$

Usually data packet is much longer than ACK frame, thus consumes more energy. The idle listening energy consumption is largely minimized due to the use of an efficient MAC protocol and low-power radio triggered wakeup. So from Eq. (12), we can know that, the majority of the energy consumption comes from data transmission and reception.

Node that, although the energy consumption is based on relatively ideal wakeup and coordination mechanisms, our framework and methodology presented in this paper should

be general enough to be applied to other coordination or wakeup/sleep mechanisms.

#### 4.4 Opportunistic energy efficiency

We propose a new local metric, opportunistic energy efficiency (OEE), which aims to strike a good balance between the routing efficiency and energy consumption. Given a forwarding candidate set  $\mathcal{F}$ , the new metric is denoted as  $OEE(\mathcal{F})$  and defined in Eq. (13).

$$OEE(\mathcal{F}) = \frac{EPA(\mathcal{F})}{P_{\mathcal{F}} \bar{E}_{or}} \cdot L_{data} \quad (13)$$

where  $L_{data}$  is the data packet length in bits. Note that  $\frac{EPA(\mathcal{F})}{P_{\mathcal{F}}}$  is the expected packet advancement for a successful opportunistic forwarding considering retransmissions. If the unit of  $EPA(\mathcal{F})$  is meter, and  $\bar{E}_{or}$  is Joule, the unit of  $OEE(\mathcal{F})$  is bmpJ. The physical meaning of  $OEE(\mathcal{F})$  is the expected bit advancement towards the destination by consuming one Joule of energy for a successful one-hop opportunistic forwarding.

#### 4.5 Problem definition

In this paper, our goal is to find  $\mathcal{F}^*$  to maximize the metric  $OEE(\mathcal{F})$  in each local opportunistic packet forwarding, which can be formulated as the following optimization problem:

$$\mathcal{F}^* = \operatorname{argmax}_{\forall \mathcal{F} \subseteq \mathcal{C}} OEE(\mathcal{F}) \quad (14)$$

The intuition to maximize the OEE is as follows. Maximizing OEE implies a large EPA in the numerator, which will result in fewer transmissions and hops for delivering a packet from the source to the destination. Maximizing OEE also implies a small denominator, which indicates lower energy consumption at each hop. Therefore maximizing OEE tends to minimize the energy consumption to forward a packet from the source to the destination, that is, to maximize the energy efficiency. Note that maximizing OEE cannot guarantee a global optimum of energy efficiency due to the lack of the global information. It is common for all the location based routing schemes that make greedy local routing decisions and trade significantly reduced routing (i.e., route establishment, update, and maintenance) overhead with suboptimal routing performance. However, we will show in Sect. 7 that our energy-efficient GOR applying OEE metric achieves better performance than blind GOR and GR solutions.

Solving this optimization problem needs to answer the following two questions: (a) How many and which nodes should be involved in the local forwarding? b) What priority should they follow to forward a packet?

This optimization problem is a combinatorial problem. One naive way to solve this problem is to enumerate all the possible candidate set  $\mathcal{F}$ , which yields a complexity of  $O(eN!)$ , where  $e$  is the base of natural logarithm and  $N$  is the number of available next-hop neighbors. It is, however, infeasible when  $N$  is large. In the next section, we will reveal some important properties about GOR on selecting and prioritizing the forwarding candidates in order to maximize the EPA, which will guide us to design a heuristic algorithm to solve the optimization problem in polynomial time.

## 5 Properties about GOR and candidate set

### 5.1 Increasing property of EPA

Intuitively, increasing the number of forwarding candidates would result in a better chance of larger packet advancement towards the destination. We first present Proposition 5.1 to confirm this intuition.

**Proposition 5.1** (Strictly increasing property) *Given  $\mathcal{C}$  with  $N$  nodes, let  $EM(\mathcal{C}, r)$  denote the maximum EPA defined in Eq. (3) achieved by selecting  $r$  nodes from  $\mathcal{C}$  as forwarding candidates, then  $EM(\mathcal{C}, r)$  is a strictly increasing function of  $r$ . That is*

$$EM(\mathcal{C}, m) < EM(\mathcal{C}, n), 1 \leq m < n \leq N \quad (15)$$

*Proof* This proposition is intuitive. Due to space limit, we omit the proof here. Interested readers can refer to [31] for a detailed proof.  $\square$

Proposition 5.1 basically indicates that the more nodes get involved in GOR, the larger the EPA can be. So the GOR that involves all the nodes in  $\mathcal{C}$  will achieve the largest EPA. However, it is not always the most energy efficient way to forward packets by involving all the nodes in  $\mathcal{C}$ . As from Eq. (12), we know one transmission is accompanied with  $r$  receptions at the forwarding candidates, so involving all the nodes in  $\mathcal{C}$  consumes the most energy. On the other hand, traditional geographic routing involving only one forwarding candidate has the least energy cost consisting of only one transmission and one reception, but it achieves the least EPA per hop, which indicates lower routing efficiency as more hops (transmissions) are needed to reach the final destination.

### 5.2 Relay priority rule

Proposition 5.2 answers the relay priority question.

**Proposition 5.2** (Relay priority rule) *Given a forwarding candidate set  $\mathcal{F}$  with  $r$  ( $r \geq 1$ ) nodes, the  $EM(\mathcal{F}, |\mathcal{F}|)$*

(the maximum EPA by using all the nodes in  $\mathcal{F}$ ) can only be obtained by giving the node closer to the destination higher relay priority, that is,  $\forall i_p, i_q \in \mathcal{F}$ , if  $d_p \geq d_q$ ,  $i_p$  has higher priority than  $i_q$ . We can break the tie by assigning arbitrary priority order to nodes with equal advancement. As  $d_1 \geq d_2 \geq \dots \geq d_r$ ,  $EM(\mathcal{F}, |\mathcal{F}|)$  can be calculated as follows.

$$EM(\mathcal{F}, |\mathcal{F}|) = \sum_{k=1}^r d_k p_k \cdot \prod_{n=0}^{k-1} (1 - p_n) \tag{16}$$

where  $p_0 := 0$ .

*Proof* We proof Proposition 5.2 by induction.  $\square$

First, for  $r = 1$ , there is only one node in  $\mathcal{F}$ , then obviously Eq. (16) holds.

Next, we assume Eq. (16) holds for  $r = N$  ( $N \geq 1$ ), we want to prove Eq. (16) holds for  $r = N + 1$ .

Assume  $|\mathcal{F}| = N + 1$ .  $\mathcal{F}$  can be divided into two complementary sub-sets,  $\mathcal{F}_1 = \mathcal{F} \setminus \{i_m\}$  with  $N$  nodes and  $\mathcal{F}_2 = \{i_m\}$  with one node. Then the EPA of any permutation  $\pi_j(\mathcal{F})$  of the  $N + 1$  nodes, can be expressed as the combination of the EPAs of the corresponding permutations of  $N$  nodes in  $\mathcal{F}_1$  and the other node  $i_m$  in  $\mathcal{F}_2$ . Then  $EM(\mathcal{F}, |\mathcal{F}|)$  is expressed as:

$$EM(\mathcal{F}, |\mathcal{F}|) = \max_{1 \leq m \leq N+1} \left\{ \sum_{k=1}^{m-1} d_k p_k \cdot \prod_{w=0}^{k-1} (1 - p_w) + \sum_{k=m+1}^{N+1} d_k p_k \cdot \frac{\prod_{w=0}^{k-1} (1 - p_w)}{1 - p_m} + d_m p_m \cdot \frac{\prod_{w=0}^{N+1} (1 - p_w)}{1 - p_m} \right\}$$

The equality holds for the definitions of  $EM(\mathcal{F}, |\mathcal{F}|)$ ,  $\mathcal{F}_1$  and  $\mathcal{F}_2$ , and the inductive hypothesis. Then we only need to prove for any integer  $m$  ( $1 \leq m \leq N$ ),

$$A := \sum_{k=1}^{m-1} d_k p_k \cdot \prod_{w=0}^{k-1} (1 - p_w) + \sum_{k=m+1}^{N+1} d_k p_k \cdot \frac{\prod_{w=0}^{k-1} (1 - p_w)}{1 - p_m} + d_m p_m \cdot \frac{\prod_{w=0}^{N+1} (1 - p_w)}{1 - p_m} \leq B := \sum_{k=1}^{N+1} d_k p_k \cdot \prod_{w=0}^{k-1} (1 - p_w)$$

Subtracting A from B, we have

$$B - A = \frac{1}{1 - p_m} \sum_{k=m+1}^{N+1} (d_m - d_k) p_m p_k \prod_{w=0}^{k-1} (1 - p_w) \geq 0$$

The above inequality holds, because  $d_m \geq d_k$ . Then the Eq. (16) holds for  $r = N + 1$ . So it holds for any  $r$  ( $r \geq 1$ ).

Proposition 5.2 indicates that when a forwarding candidate set is given, the maximum EPA can be achieved by assigning the relay priority to each node based on their

distances to the destination. That is, the furthest node should try to forward the packet first; if it failed (i.e., did not receive the packet correctly), the second furthest node should try next, and so on. For example, in Fig. 1, if nodes  $i_2$  and  $i_4$  are chosen as forwarding candidates, we should assign  $i_2$  higher reliability than  $i_4$ . If two nodes have the same advancement towards the destination, we can assign either one with higher priority. We will show in Sect. 6, considering reliability, we will choose the candidates who contributes more on the reliability if they have the same expected advancement.

Note that this proposition does not mean we should always select  $i_1$  and  $i_2$  in Fig. 1 in order to maximize the EPA when the candidate set size is 2. Actually, choosing  $i_2$  and  $i_3$  maximizes the EPA when two forwarding candidates are involved.

A coordination scheme similar to the ones in [3, 4, 25] is necessary here to coordinate the transmission among the forwarding candidates. However, the detailed design of the coordination scheme is beyond the scope of this paper. In this paper, we assume the relay priority can be ideally followed. The analysis result is the upper bound of the EPA that GOR can achieve.

Based on Proposition 5.2, we redefine  $EM(\mathcal{C}, r)$  as follows,

$$EM(\mathcal{C}, r) = \max_{1 \leq j \leq C_N^r} \left\{ \sum_{k=1}^r d_{j_k} p_{j_k} \cdot \prod_{n=0}^{k-1} (1 - p_{j_n}) \right\} \tag{17}$$

where  $N = |\mathcal{C}|$ , and  $C_N^r$  is the number of combinations of selecting  $r$  ( $1 \leq r \leq N$ ) nodes from the  $N$  nodes in  $\mathcal{C}$ . For each combination  $j$ , we have  $d_{j_1} > d_{j_2} > \dots > d_{j_r}$  and  $p_{j_0} := 0$ . Here the number of choices is the number of combinations, as when  $r$  nodes are selected, the order of them is already determined by the *Relay priority rule* (Proposition 5.2).

Next, we will identify and prove two important properties about the  $EM(\mathcal{C}, r)$  function. First, we look at the characteristics of the forwarding candidates that are selected to achieve  $EM(\mathcal{C}, r)$  with various sizes  $r$ . We prove the *Containing property* for those node sets. Following that, the *Concavity* of the function  $EM(\mathcal{C}, r)$  is proved.

### 5.3 Containing property of feasible candidate set

Define  $\mathcal{F}_r^*$  as a feasible forwarding candidate set that achieves the  $EM(\mathcal{C}, r)$ . For a given  $\mathcal{C}$ , we have the following containing property of  $\mathcal{F}_r^*$ 's.

**Lemma 5.3** (Containing property) *Given the available next-hop node set  $\mathcal{C}$  with  $N(1 \leq N < \infty)$  nodes,  $\forall \mathcal{F}_{r-1}^*, \exists \mathcal{F}_r^*$ , s.t.*

$$\mathcal{F}_{r-1}^* \subset \mathcal{F}_r^* \quad \forall 1 \leq r \leq N \tag{18}$$

*Proof of Lemma 5.3* Please refer to “Appendix 1” for the proof.  $\square$

Lemma 5.3 indicates that an  $(r - 1)$ -node set that achieves  $EM(\mathcal{C}, r - 1)$  is a subset of one feasible  $r$ -node set that achieves  $EM(\mathcal{C}, r)$ .

#### 5.4 Concavity of maximum EPA

Following Lemma 5.3, we have the concave property of  $EM(\mathcal{C}, r)$  as in Proposition 5.4.

**Proposition 5.4** (Concavity of maximum EPA) *Given the available next-hop node set  $\mathcal{C}$  with  $N$  ( $N \geq 1$ ) nodes,  $EM(\mathcal{C}, r)$ , defined in Eq. (17), is a concave function of  $r$ . That is,  $\forall 1 \leq r < N, EM(\mathcal{C}, r + 1) - EM(\mathcal{C}, r) < EM(\mathcal{C}, r) - EM(\mathcal{C}, r - 1)$ .*

*Proof* According to Lemma 5.3, assume  $\mathcal{F}_{r+1}^* \setminus \mathcal{F}_r^* = \{i_k\}$ , and  $\mathcal{F}_r^* \setminus \mathcal{F}_{r-1}^* = \{i_j\}$ . There are two cases for the advancement relationship between node  $i_k$  and  $i_j$ .

(1)  $d_k \geq d_j$ .

Then  $\mathcal{F}_{r+1}^*, \mathcal{F}_r^*$  and  $\mathcal{F}_{r-1}^*$  can be represented as

$$\begin{aligned} \mathcal{F}_{r+1}^* &= \langle \mathcal{A}_1, i_k, \mathcal{A}_2, i_j, \mathcal{A}_3 \rangle, \\ \mathcal{F}_r^* &= \langle \mathcal{A}_1, \mathcal{A}_2, i_j, \mathcal{A}_3 \rangle, \\ \mathcal{F}_{r-1}^* &= \langle \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \rangle \end{aligned} \tag{19}$$

where  $\mathcal{A}_i (1 \leq i \leq 3)$  is ordered node set and can be  $\emptyset$ . Here, we also handle  $\mathcal{A}_i$  as one node by giving the same definition of  $d_{\mathcal{A}_i}$  and  $p_{\mathcal{A}_i}$  as in Eq. (35) (see “Appendix 1”). The following advancement relationship holds that

$$d_{\mathcal{A}_1} \geq d_k \geq d_{\mathcal{A}_2} \geq d_j \geq d_{\mathcal{A}_3} \tag{20}$$

Because  $EPA(\mathcal{F}_r^*)$  is the largest EPA achieved by selecting  $r$  nodes, we have

$$\begin{aligned} B &:= EPA(\mathcal{F}_r^*) - EPA(\langle \mathcal{A}_1, i_k, \mathcal{A}_2, \mathcal{A}_3 \rangle) \\ &= (1 - p_{\mathcal{A}_1})[p_k \cdot EPA(\mathcal{A}_2) - EPA(i_k)] \\ &\quad + (1 - p_{\mathcal{A}_1})(1 - p_{\mathcal{A}_2})[(p_k - p_j)EPA(\mathcal{A}_3) + EPA(i_j)] \geq 0 \end{aligned} \tag{21}$$

Then,

$$\begin{aligned} [EPA(\mathcal{F}_r^*) - EPA(\mathcal{F}_{r-1}^*)] - [EPA(\mathcal{F}_{r+1}^*) - EPA(\mathcal{F}_r^*)] \\ = B + (1 - p_{\mathcal{A}_1})(1 - p_{\mathcal{A}_2})p_k p_j (d_j - EPA(\mathcal{A}_3)) > 0 \end{aligned} \tag{22}$$

Inequality (22) holds because  $B \geq 0$  (inequality (21)) and  $d_j - EPA(\mathcal{A}_3) > d_j - d_{\mathcal{A}_3} \geq 0$ .

So  $EPA(\mathcal{F}_r^*) - EPA(\mathcal{F}_{r-1}^*) > EPA(\mathcal{F}_{r+1}^*) - EPA(\mathcal{F}_r^*)$  when  $d_k \geq d_j$ .

(2)  $d_k < d_j$ .

Then  $\mathcal{F}_{r+1}^*, \mathcal{F}_r^*$  and  $\mathcal{F}_{r-1}^*$  can be represented as

$$\begin{aligned} \mathcal{F}_{r+1}^* &= \langle \mathcal{A}_1, i_j, \mathcal{A}_2, i_k, \mathcal{A}_3 \rangle, \\ \mathcal{F}_r^* &= \langle \mathcal{A}_1, i_j, \mathcal{A}_2, \mathcal{A}_3 \rangle, \\ \mathcal{F}_{r-1}^* &= \langle \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \rangle \end{aligned} \tag{23}$$

Similarly, with

$$B := EPA(\mathcal{F}_r^*) - EPA(\langle \mathcal{A}_1, \mathcal{A}_2, i_k, \mathcal{A}_3 \rangle) \geq 0 \tag{24}$$

we can derive that

$$\begin{aligned} [EPA(\mathcal{F}_r^*) - EPA(\mathcal{F}_{r-1}^*)] - [EPA(\mathcal{F}_{r+1}^*) - EPA(\mathcal{F}_r^*)] \\ = B + (1 - p_{\mathcal{A}_1})(1 - p_{\mathcal{A}_2})p_k p_j (d_k - EPA(\mathcal{A}_3)) > 0 \end{aligned} \tag{25}$$

So  $EPA(\mathcal{F}_r^*) - EPA(\mathcal{F}_{r-1}^*) > EPA(\mathcal{F}_{r+1}^*) - EPA(\mathcal{F}_r^*)$  when  $d_k < d_j$ .

From the analysis above, we know for arbitrary finite number  $N$ , we have  $EPA(\mathcal{F}_r^*) - EPA(\mathcal{F}_{r-1}^*) > EPA(\mathcal{F}_{r+1}^*) - EPA(\mathcal{F}_r^*), \forall 1 \leq r \leq N - 1$ , that is,  $EM(\mathcal{C}, r)$  defined in Eq. (17) is a concave function of  $r$ .  $\square$

Combining Propositions 5.1 and 5.4, we know that giving an available next-hop node set  $\mathcal{C}$  of node  $i$  with  $N$  nodes, the maximum EPA of selecting  $r$  ( $1 \leq r \leq N$ ) nodes is a *strictly increasing and concave function* of  $r$ . This means that although the maximum EPA keeps increasing when more nodes get involved, the speed of the increase slows down. When more nodes are involved, the gained extra EPA becomes smaller and smaller.

#### 5.5 Reliability increasing property

Following the *Containing property* in Lemma 5.3, we have the *Reliability increasing property* in Corollary 5.5.

Denote  $\mathcal{F}_r^* = \langle i_{r_1}, i_{r_2}, \dots, i_{r_r} \rangle$ . Define the one-hop reliability  $P_{\mathcal{F}_r^*}$  in Eq. (26) which is the probability of at least one node in  $\mathcal{F}_r^*$  correctly receiving the packet sent by node  $i$  for one transmission.

$$P_{\mathcal{F}_r^*} = 1 - \prod_{n=0}^r (1 - p_{r_n}) \tag{26}$$

where  $p_{r_0} := 0$ .

Define  $P^*(r)$  in Eq. (27) which is the maximum one-hop reliability achieved by one of the feasible  $\mathcal{F}_r^*$ 's.

$$P^*(r) = \max_{\forall \mathcal{F}_r^*} \{P_{\mathcal{F}_r^*}\} \tag{27}$$

**Corollary 5.5**  $P^*(r)$  defined in Eq. (27) is an increasing function of  $r$ .

*Proof* The proof is straightforward following Lemma 5.3. Assume one  $\mathcal{F}_r^*$  achieves  $P^*(r)$ , then  $\exists \mathcal{F}_{r+1}^*$  s.t.



$\mathcal{F}_{r+1}^* \supset \mathcal{F}_r^*$ . According to the definitions of  $P^*(r)$  in Eq. (27) and the one-hop reliability in Eq. (26), we have

$$P^*(r+1) \geq P_{\mathcal{F}_{r+1}^*} > P_{\mathcal{F}_r^*} = P^*(r)$$

So  $P^*(r)$  is an increasing function of  $r$ . □

Corollary 5.5 indicates that the maximum one-hop reliability corresponding to the forwarding candidate set that maximizes the EPA also increases when more forwarding candidates are involved. The increasing of the maximum EPA implies increasing of the reliability. Therefore, the EPA is a good metric for balancing the packet advancement and reliability.

### 6 Selecting forwarding candidates in polynomial time

Based on the properties proved in Sect. 5, we are ready to propose two heuristic polynomial time algorithms to find an approximation of the optimal forwarding candidate set  $\mathcal{F}^*$  in Eq. (14). One algorithm is based on Proposition 5.2 and Lemma 5.3, and the other one is a dynamic programming based on Proposition 5.2.

#### 6.1 Algorithm based on Lemma 5.3

The input of Algorithm 1 is the next-hop neighboring node set ( $\mathcal{C}$ ) and the function to compute  $\bar{E}_{or}$ . Algorithm 1 searches the optimal candidate sets  $\mathcal{F}'$  that maximizes EPA with different number of candidates. Based on the *Containing property* in Lemma 5.3, in order to find an  $\mathcal{F}'$  with  $r$  nodes, we add a new node into the optimal candidate set with  $r - 1$  nodes. For feasible sets having the same maximum EPA, we choose the one that achieves higher one-hop reliability (lines 9 and 10). During the search, we compute the OEE for each examined candidate set, and find the one ( $\mathcal{F}^\circ$ ) that maximizes the OEE ( $G^\circ$ ) in lines 12 and 13. Note that  $\mathcal{F}^\circ$  is not necessary the same as  $\mathcal{F}'$ .

It is not difficult to find an algorithm to calculate  $EPA(\mathcal{F})$  and  $G$  (in line 8) in  $\mathbf{O}(|\mathcal{F}|)$  running time. So Algorithm 1 costs  $\mathbf{O}(N^3)$  running time.

Table 1 shows the procedure of finding the  $G^\circ$  and  $\mathcal{F}^\circ$  by applying Algorithm 1 on the example in Fig. 1. We assume the following parameters:  $E_{data\_tx} = 1, E_{data\_rx} = 0.5, E_{to} = 0.025, E_{ACK\_tx} = 0.1, E_{ACK\_rx} = 0.05, E_{SIFS} = 0$ .  $L_{data}$  is normalized as 1 unit. The procedure runs from Round 1 to Round 5, and in each round it runs from the top to the bottom. In the first round,  $\langle i_3 \rangle$  is found as the node achieving the maximum EPA as well as the maximum OEE by selecting one forwarding candidate; in the second round,

#### Algorithm 1 Finding an $\mathcal{F}^\circ$ based on Lemma 5.3

---

```

1: Input:  $\mathcal{C}, \bar{E}_{or}(\cdot)$ 
2: {Initialization}
3:  $G^\circ \leftarrow 0; A' \leftarrow A'_c \leftarrow 0;$ 
4:  $\mathcal{F}' \leftarrow \mathcal{F}'_c \leftarrow \emptyset; P_{\mathcal{F}'} \leftarrow P_{\mathcal{F}'_c} \leftarrow 0; \mathcal{B} \leftarrow \mathcal{C};$ 
5: { $\mathcal{B}$  is the remained node set}
6: while  $\mathcal{B} \neq \emptyset$  do
7:   for all  $i_j \in \mathcal{B}$  do
8:      $\mathcal{F} \leftarrow \mathcal{F}' \cup \{i_j\}; P_{\mathcal{F}} \leftarrow 1 - (1 - P_{\mathcal{F}'})(1 - p_j); A \leftarrow EPA(\mathcal{F});$ 
        $G = A / (P_{\mathcal{F}} \cdot \bar{E}_{or}(\mathcal{F}));$ 
9:     if  $A > A'_c \parallel (A = A'_c \ \& \ P_{\mathcal{F}} > P_{\mathcal{F}'_c})$  then
10:        $A'_c \leftarrow A; \mathcal{F}'_c \leftarrow \mathcal{F}; P_{\mathcal{F}'_c} \leftarrow P_{\mathcal{F}};$ 
11:     end if
12:     if  $G > G^\circ$  then
13:        $G^\circ = G; \mathcal{F}^\circ \leftarrow \mathcal{F};$ 
14:     end if
15:   end for
16:    $\mathcal{B} \leftarrow \mathcal{C} \setminus \mathcal{F}'_c; A' \leftarrow A'_c; \mathcal{F}' \leftarrow \mathcal{F}'_c; P_{\mathcal{F}'} \leftarrow P_{\mathcal{F}'_c};$ 
17: end while
18: return  $(G^\circ, \mathcal{F}^\circ)$ 

```

---

we test the sets containing  $i_3$ , and  $\langle i_2, i_3 \rangle$  is found as the optimal node set by selecting two forwarding candidates; in the third round, we test the set containing  $i_2$  and  $i_3$ , and  $\langle i_1, i_2, i_3 \rangle$  is found as the optimal node set by selecting three forwarding candidates. We can find that although the EPA increases, OEE drops. Since by involving more forwarding candidates, the advancement gain is marginal, while more energy is consumed at the candidate side. Finally, we find that  $\langle i_2, i_3 \rangle$  achieves the maximum OEE.

We further use a brute-force algorithm to enumerate all the possible forwarding candidate sets and find that the optimal  $\mathcal{F}^*$  that achieves the maximum OEE in Eq. (14) is also  $\langle i_2, i_3 \rangle$ , which indicates that Algorithm 1 indeed finds the optimal forwarding candidate set in this example.

#### 6.2 Dynamic programming algorithm

We now introduce another efficient dynamic programming algorithm.

Recall that nodes  $i_j$ 's ( $1 \leq j \leq N$ ) in  $\mathcal{C}$  are ordered according to the advancements as  $d_1 \geq d_2 \geq \dots \geq d_N$ . Denote the set  $\langle i_q, i_{q+1}, \dots, i_N \rangle$  ( $1 \leq q \leq N$ ) as  $\mathcal{C}_q$ . Following the denoting,  $\mathcal{C}_1 = \mathcal{C}$ . According to the *Relay priority rule* in Proposition 5.2 and the definition of  $EM(\mathcal{C}, r)$ , we then have,

$$EM(C_q, r) = \begin{cases} 0, & r = 0 \quad \text{or} \quad N - q + 1 < r; \\ \text{Max}\{d_q p_q + (1 - p_q)EM(C_{q+1}, r - 1), \\ EM(C_{q+1}, r)\}, & \text{Otherwise.} \end{cases} \quad (28)$$

**Table 1** Procedure of finding a maximum energy efficiency *OEE* and the corresponding forwarding candidate set  $\mathcal{F}^\circ$  using Algorithm 1 on the example in Fig. 1

Round 1			Round 2			Round 3			Round 4			Round 5		
$\mathcal{F}$	EPA	OEE	$\mathcal{F}$	EPA	OEE	$\mathcal{F}$	EPA	OEE	$\mathcal{F}$	EPA	OEE	$\mathcal{F}$	EPA	OEE
$\langle i_1 \rangle$	3.75	2.38	$\langle i_1, i_3 \rangle$	11.4	5.17	$\langle i_1, i_2, i_3 \rangle$	15.14	5.41	$\langle i_1, i_2, i_3, i_4 \rangle$	16.57	4.86	$\langle i_1, i_2, i_3, i_4, i_5 \rangle$	16.84	4.21
$\langle i_2 \rangle$	8	4.97	$\langle i_2, i_3 \rangle$	13.4	6.02	$\langle i_2, i_3, i_4 \rangle$	15.08	5.32	$\langle i_1, i_2, i_3, i_5 \rangle$	16.06	4.70			
$\langle i_3 \rangle$	9	5.49	$\langle i_3, i_4 \rangle$	11.8	5.24	$\langle i_2, i_3, i_5 \rangle$	14.48	5.09						
$\langle i_4 \rangle$	7	4.23	$\langle i_3, i_5 \rangle$	10.8	4.76									
$\langle i_5 \rangle$	4.5	2.67												

$E_{data\_tx} = 1, E_{data\_rx} = 0.5, E_{to} = 0.025, E_{ACK\_tx} = 0.1, E_{ACK\_rx} = 0.05, E_{SIFS} = 0$

$EM(C, r)$  ( $1 \leq r \leq N$ ) can be efficiently calculated by applying Eq. (28) recursively using dynamic programming [32].

The pseudocode of the dynamic programming algorithm GetM-B is given in Algorithm 2, where  $|\mathcal{C}| = N, F_{(q,r)}$  is the ordered node set corresponding to  $EM(C_q, r), P_{(q,r)}$  is the corresponding one-hop reliability, and  $d_i$ 's are sorted in descending order ( $d_1 \geq d_2 \geq \dots \geq d_N$ ). We also choose the feasible set that achieves higher one-hop reliability when two feasible sets have the same EPA (lines 16 to 17). The complexity of Algorithm 2 is also  $\mathbf{O}(N^3)$ .

Table 2 shows the procedure of finding the  $G^\circ$  and  $\mathcal{F}^\circ$  by applying Algorithm 2 on the example in Fig. 1. We use the same settings as in Table 1. The procedure runs from Round 1 to Round 5, and in each round it runs from the bottom to the top. Note that although it finds the same  $\mathcal{F}^\circ = \langle i_2, i_3 \rangle$  as in Table 1, some of the tested node sets are different from the ones in Table 1. This  $\mathcal{F}^\circ$  is also indeed the  $\mathcal{F}^*$  achieving the maximum OEE in Eq. (14).

### 6.3 Algorithms 1 and 2 find optimal set

We exam how close the OEE and  $\mathcal{F}^\circ$  obtained from Algorithms 1 and 2 are to the optimal ones. We randomly generate  $(d, p)$  pairs under different nexthop available neighbor size ( $N$ ).  $d$  is within [1, 10] and  $p$  is within (0.1,1).

$E_{data\_tx} = 1, E_{data\_rx} = 0.5, E_{to} = 0.025,$   
 $E_{ACK\_tx} = 0.1, E_{ACK\_rx} = 0.05, E_{SIFS} = 0.$

We find that for each  $N$  from 1 to 8, Algorithms 1 and 2 always find the optimal set  $\mathcal{F}^*$  in 1,000 random tests.

### Algorithm 2 Finding an $\mathcal{F}^\circ$ by dynamic programming

```

1: Input:  $C, \bar{E}_{or}(\cdot)$ 
2: {Initialization}
3: for  $i \leftarrow 1$  to  $N$  do
4:    $EM(C_{i+1}, 0) \leftarrow 0; F_{(i+1,0)} \leftarrow \emptyset; P_{(i+1,0)} \leftarrow 0;$ 
5: end for
6:  $G^\circ \leftarrow 0; \mathcal{F}^\circ \leftarrow \emptyset; EM(C_{N+1}, 1) \leftarrow 0; P_{(N+1,1)} \leftarrow 0;$ 
7: for  $i \leftarrow 1$  to  $N$  do
8:   for  $q \leftarrow N-r+1$  down to 1 do
9:      $A \leftarrow d_q p_q + (1 - p_q)EM(C_{q+1}, r - 1);$ 
10:     $P \leftarrow 1 - (1 - P_{(q+1,r-1)})(1 - p_q);$ 
11:     $\mathcal{F} = \{i_q\} \cup F_{(q+1,r-1)};$ 
12:     $G = A/P/\bar{E}_{or}(\mathcal{F});$ 
13:    if  $G > G^\circ$  then
14:       $G^\circ = G; \mathcal{F}^\circ = \mathcal{F};$ 
15:    end if
16:    if  $A > EM(C_{q+1}, r) || (A == EM(C_{q+1}, r) \ \&\& \ P > P_{(q+1,r)})$ 
    then
17:       $EM(C_q, r) \leftarrow A;$ 
18:       $F_{(q,r)} \leftarrow F_{(q+1,r-1)} \cup \{i_q\}; P_{(q,r)} \leftarrow P;$ 
19:    else
20:       $EM(C_q, r) \leftarrow EM(C_{q+1}, r); F_{(q,r)} \leftarrow F_{(q+1,r)};$ 
21:       $P_{(q,r)} \leftarrow P_{(q+1,r)};$ 
22:    end if
23:  end for
24: end for
25: return  $(G^\circ, \mathcal{F}^\circ);$ 

```

## 7 Performance evaluation

We evaluate the performance of our energy-efficient GOR (EGOR) scheme applying the candidate selection

**Table 2** Procedure of finding a maximum energy efficiency  $OEE$  and the corresponding forwarding candidate set  $\mathcal{F}^\circ$  using Algorithm 2 on the example in Fig. 1

$q$	Round 1		Round 2		Round 3		Round 4		Round 5	
	$F_{(q,1)}$	$EM(C_q, 1)$	$F_{(q,2)}$	$EM(C_q, 2)$	$F_{(q,3)}$	$EM(C_q, 3)$	$F_{(q,4)}$	$EM(C_q, 4)$	$F_{(q,5)}$	$EM(C_q, 5)$
1	$\langle i_3 \rangle$	9	$\langle i_2, i_3 \rangle$	13.4	$\langle i_1, i_2, i_3 \rangle$	15.14	$\langle i_1, i_2, i_3, i_4 \rangle$	16.57	$\langle i_1, i_2, i_3, i_4, i_5 \rangle$	16.84
2	$\langle i_3 \rangle$	9	$\langle i_2, i_3 \rangle$	13.4	$\langle i_2, i_3, i_4 \rangle$	15.08	$\langle i_2, i_3, i_4, i_5 \rangle$	15.40		
3	$\langle i_3 \rangle$	9	$\langle i_3, i_4 \rangle$	11.8	$\langle i_3, i_4, i_5 \rangle$	12.34				
4	$\langle i_4 \rangle$	7	$\langle i_4, i_5 \rangle$	8.35						
5	$\langle i_5 \rangle$	4.5								
	$G^\circ = 5.49, \mathcal{F}^\circ = \langle i_3 \rangle$		$G^\circ = 6.02, \mathcal{F}^\circ = \langle i_2, i_3 \rangle$		$G^\circ = 6.02, \mathcal{F}^\circ = \langle i_2, i_3 \rangle$		$G^\circ = 6.02, \mathcal{F}^\circ = \langle i_2, i_3 \rangle$		$G^\circ = 6.02, \mathcal{F}^\circ = \langle i_2, i_3 \rangle$	

$E_{data\_tx} = 1, E_{data\_rx} = 0.5, E_{to} = 0.025, E_{ACK\_tx} = 0.1, E_{ACK\_rx} = 0.05, E_{SIFS} = 0$

algorithms proposed in Sect. 6 under a sensor network scenario. We also compare the performance of EGOR with geographic routing (denoted as ‘‘GR’’) which uses only one forwarding candidate that achieves the maximum EPA, and blind geographic opportunistic routing (denoted as ‘‘Blind GOR’’) which involves all the available next-hop nodes as forwarding candidates. Recall that we focus ourselves on greedy forwarding. That is, when the packet gets stuck due to a communication void where no available next-hop neighbors are closer to the destination than the current node, the packet will be dropped in all the three schemes.

Without explicit explanation, we assume an ideal coordination and low-power triggered wakeup schemes are used, i.e. the forwarding priority is strictly followed and the idle listening energy consumption is minimized.

### 7.1 Simulation setup

The simulated sensor network has stationary nodes uniformly distributed in a  $100 \times 100 m^2$  square region, with nodes having identical fixed transmission power of 5dbm. The source and the sink nodes are fixed at two corners across the diagonal of the square area. Various situations are examined by varying node densities, retransmission limits, and packet sizes. For each setting, we perform 50 independent runs, with 3,000 packets sent in each run. Node locations are randomly re-assigned and PRRs between nodes are re-calculated in each run. Each result in the figures represents the average value of the 50 runs, with 95% confidence interval indicated.

*Channel Model:* To simulate a realistic channel model for lossy WSNs, we use the log-normal shadowing path loss model derived in [3]. This model considers several environmental and radio parameters<sup>2</sup>, such as the path-loss

<sup>2</sup> Please refer to [3] for a complete description of the model.

exponent ( $\alpha$ ) and log-normal shadowing variance ( $\sigma$ ) of the environment, and the modulation and encoding schemes of the radio. Eq. (29) calculates the packet reception ratio as a function of the packet size, Signal to Noise Ratio (SNR), and distance between a transmitter and receiver. It resembles a MICA2 mote [33], which has data rate of 19.2 kbps, and the noise bandwidth 30 kHz. Non-coherent FSK and Manchester are used as the modulation and encoding schemes ( $\rho = 2$ ), respectively. The environmental parameters are set to  $\alpha = 3.5$  and  $\sigma = 4$ . In reality, the SNR changes according to the variance of path-loss due to shadowing. To simulate a long-term (in seconds) PRR on each link, we average the PRR under 1000 channel realizations (SNRs) of the log-normal shadowing model. We assume the long-term PRR on each link is stable in each run.

$$PRR(L_p, d) = \left(1 - \frac{1}{2}e^{-\frac{\gamma(d)}{2 \times 0.64}}\right)^{8L_{pr}} \left(1 - \frac{1}{2}e^{-\frac{\gamma(d)}{2 \times 0.64}}\right)^{8\rho(L_p - L_{pr})} \tag{29}$$

where  $d$  is the transmitter-receiver distance,  $\gamma(d)$  is the SNR,  $\rho$  is the encoding ratio,  $L_p$  is the packet (including preamble, header, and payload) length in bytes, and  $L_{pr}$  is the physical layer preamble length in bytes.  $L_{pr}$  is fixed at 20 bytes to resemble a MICA2 mote in the simulation.

*Energy Consumption Model:* The energy consumption is obtained by multiplying the power consumption and the packet transmission/reception time as in Eq. (30). Under 5dbm transmission power, according to the power consumption for the MICA2 [33], the transmission power consumption ( $P_{tx}$ ) is 81mw, and both reception power consumption ( $P_{rx}$ ) and ( $P_{idle}$ ) are 30mw.

$$E_{data\_tx} = \frac{8L_p P_{tx}}{R}, E_{data\_rx} = \frac{8L_p P_{rx}}{R},$$

$$E_{ACK\_tx} = \frac{8(L_{pr} + L_h)P_{tx}}{R}, E_{ACK\_rx} = \frac{8(L_{pr} + L_h)P_{rx}}{R},$$

$$E_{to} = P_{idle}t_{to}, E_{SIFS} \approx 0 \tag{30}$$

where  $R$  is the data rate, which is fixed at 19.2kbps in our simulation. The  $L_h$  is the length of the MAC header containing the forwarding candidates' addresses. A neighbor of a transmitter has to read these addresses to decide if it is a forwarding candidate or destination of this packet.  $L_h$  is various depending on the number of forwarding candidates. Assuming each MAC address length is 2 bytes,  $L_h = 2r$  bytes, where  $r$  is the number of forwarding candidates. We assume  $t_{io}$  is 100  $\mu$ s, which is actually much shorter than the transmission time of data and ACK frames.

**Evaluation Metrics:** We define the following metrics to evaluate the performance of these three routing schemes.

- **Packet delivery ratio (PDR):** percentage of packets sent by the source that actually reach the sink. This is a measure for reliability.
- **Energy efficiency  $\eta(S, D)$ :** this metric is measured in bit-meters per Joule (bmpJ). It is calculated as in Eq. (31),

$$\eta(S, D) = \frac{8 \cdot L_p \cdot N_s \cdot PDR \cdot Dist(S, D)}{E_{total}} \quad (31)$$

where  $N_s$  denotes the number of packets sent out from source  $S$ ,  $L_p$  is the packet length in bytes, and  $E_{total}$  is the (transmission, reception, and idle listening) energy consumed by all the nodes (including those neighbors which overhear the packets but only read the header of the packets) involved in the routing procedure. We account for the distance factor, because the energy efficiency is indeed relevant to the distance between the communication pair due to the lossy property of multi-hop wireless links in WSNs.

## 7.2 Simulation results and analysis

### 7.2.1 Impact of node density

We use different node numbers (100, 125, 150, 175) to achieve various node densities corresponding to the

average number of neighbors per node of 10, 12, 15, 17, respectively. The retransmission limit is set as 7. The packet size is 75 bytes. It takes more than 10 hops to reach the destination under each density for all the three routing schemes.

Figure 2(a) shows that EGOR achieves better energy efficiency than the other two routing schemes under different node densities. For every forwarding decision, EGOR chooses the forwarding set that maximizes the OEE, the EPA per unit energy consumption. Blind GOR involving all the available next-hop nodes in the routing achieves the worst energy efficiency. The main reason is that although Blind GOR achieves the highest per transmission reliability and uses fewer transmissions to deliver all the packets, it costs much more energy on reception. The excessive reception energy cost in Blind GOR due to unnecessarily involving all the available next-hop nodes in forwarding overwhelms the benefit of reliability.

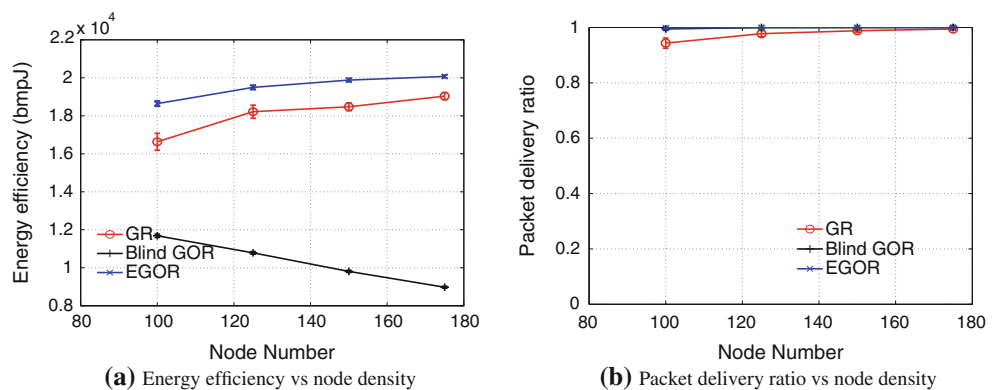
Figure 2(b) shows the PDR of these three schemes. Both EGOR and Blind GOR achieve 100% PDR under different node densities, since they involve multiple candidates into the packet forwarding and achieve high reliability. GR has the lowest reliability since it only involves one pre-defined next-hop node in each packet forwarding.

Actually, in this simulation, EGOR selects around 2.5 forwarding candidates on average under each node density. This observation suggests that only involving a few candidates is sufficient to take advantage of spatial diversity gain of OR. Blindly involving all the available next-hop nodes in GOR is an energy wasteful method.

### 7.2.2 Impact of packet size

Due to different applications and traffic types, the packet sizes can be different. We examine the impact of packet size on the performance of these three schemes. The node number is still set as 125, and the retransmission limit is 7. We set packet size at 75, 100, 125, and 150 bytes.

**Fig. 2** Simulation results under different node densities



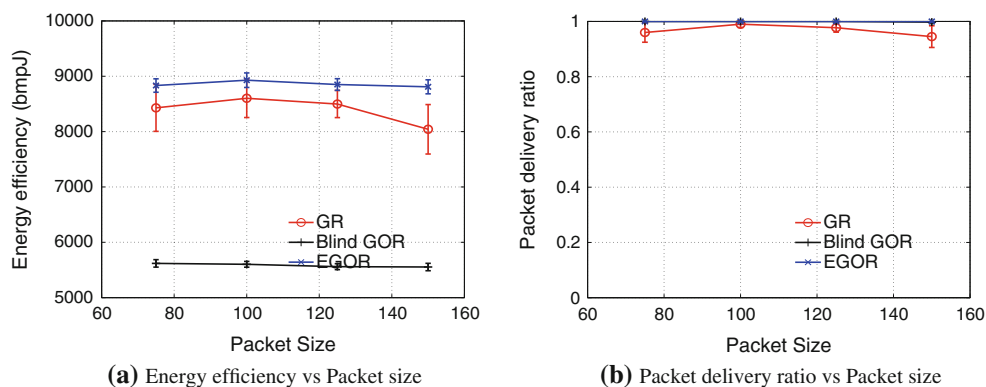
**Fig. 3** Simulation results under different packets sizes

Figure 3(a) shows that EGOR achieves better energy efficiency than the other two schemes under different packet sizes. Blind GOR achieves the lowest energy efficiency. When packet size is larger than 100, the energy efficiency of GR is decreased as packet size is increased. When the packet becomes larger, it is more likely to be corrupted by the wireless channel. The energy efficiency has the same trend as the PDR in Fig. 3(b). An interesting observation is that the PDR and energy efficiency of GR are decreased when the packet size changes from 100 to 75, under which, GR selects some longer links which have lower PRR. However, the PDR of EGOR and blind GOR remains 1 under all the packet sizes. Since they involve multiple neighbors in the packet forwarding, although the link quality is degraded when the packet size is increased, the reliability is not affected. EGOR achieves a good balance between the reliability and energy cost by smartly involving only a few “good” neighbors in the opportunistic forwarding, so it yields the best energy efficiency.

### 7.3 Impact of MAC retransmission

We have done simulation to study how the retransmission limit affects the performance of the three schemes. Due to space limit, we do not show the results, instead we make the following summary.

The benefit of increasing retransmission limit for the Blind GOR is marginal but for the GR is obvious (especially when retransmission limit is less than 4). EGOR achieves the best energy efficiency among the three schemes when the retransmission is applied. When there is no retransmission, EGOR achieves similar energy efficiency as Blind GOR. When retransmission limit is larger than 3, the energy efficiency of EGOR does not change much as the PDR is already approaching to 1.

## 8 Conclusion and future work

In this paper, we study the GOR with energy efficiency as a major concern. We introduce the definition of EPA for arbitrary number of forwarding candidates in GOR and propose a new routing metric, OEE, which evaluates EPA per unit of energy consumption for each local opportunistic forwarding. OEE jointly takes into account packet advancement, reliability, and energy consumption in GOR. Through theoretical analysis, we first show that the maximum EPA can only be achieved by following a relay priority rule – giving the forwarding candidates closer to the destination higher relay priorities when a forwarding candidate set is given. We also show that giving an available next-hop neighbor set with  $N$  nodes, the maximum EPA achieved by selecting  $r$  nodes is a strictly increasing and concave function of  $r$ . We further show that when achieving the maximum EPA, an  $r$ -node subset is contained in an  $r + 1$ -node subset. However, solely maximizing the packet advancement at each hop by involving all the available next-hop nodes in the opportunist routing cannot achieve a good energy efficiency. Aiming at striking a good balance between routing efficiency and energy consumption, by leveraging the proved findings, we propose two localized candidate selection algorithms with  $O(N^3)$  running time. The algorithms efficiently determine the forwarding candidate set that maximizes the local metric, OEE. The performance of EGOR applying the algorithms maximizing the OEE is studied through extensive simulations and compared with those of the existing geographic routing and blind opportunistic routing schemes. The results show that EGOR achieves the best energy efficiency among the three schemes under different network densities, MAC retransmission limits, and packet sizes. The energy efficiency and control packet overhead of these three routing schemes are also studied under a state-of-the-art candidate coordination scheme. Our simulation results also show that although the EPA can be maximized by involving the most number of nodes in GOR, in terms of

energy efficiency, only a very small number (around 2.5) of forwarding candidates are needed on average.

In this paper, we assume relatively ideal MAC and wakeup mechanisms in order to get more insight about the impact of candidate selection on the energy efficiency of GOR. However, the OEE proposed in this paper is general enough to be applied to different MAC protocols and wakeup/sleep schemes as long as the expected per-hop energy consumption can be formulated. We plan to test OEE in sensor network testbeds in the future.

**Acknowledgments** This work was supported in part by the US National Science Foundation under grants CNS-0746977, CNS-0716306, and CNS-0831628.

### Appendix 1

#### Proof of Lemma 5.3

*Proof* We begin by proving the following useful result. Let  $\mathcal{A}$  be an ordered node set  $\langle a_1, a_2, \dots, a_M \rangle^3$  with  $M$  ( $1 \leq M < \infty$ ) nodes, in which the nodes are ordered as  $d_{a_1} > d_{a_2} > \dots > d_{a_M}$ . Let  $\mathcal{B}$ ,  $\langle b_1, b_2, \dots, b_N \rangle$  with  $N$  ( $0 \leq N < M$ ) nodes, be a subset of  $\mathcal{A}$  such that  $d_{b_1} > d_{b_2} > \dots > d_{b_N}$  and  $b_N = a_M$ . For any node  $q \notin \mathcal{A}$  with  $d_q \leq d_{a_M}$ , we have the following inequality

$$\begin{aligned} \Delta A &:= \text{EPA}(\langle \mathcal{A}, q \rangle) - \text{EPA}(\langle q, \mathcal{A} \rangle) \\ &\geq \Delta B := \text{EPA}(\langle \mathcal{B}, q \rangle) - \text{EPA}(\langle q, \mathcal{B} \rangle) \end{aligned} \quad (32)$$

The equality holds only when  $d_{a_1} = d_{a_2} = \dots = d_{a_M} = d_q$ .

As  $\mathcal{B} \subset \mathcal{A}$ , we can represent  $\mathcal{A}$  as

$$\mathcal{A} = \langle \mathcal{A}_1, b_1, \mathcal{A}_2, b_2, \dots, \mathcal{A}_N, b_N \rangle \quad (33)$$

where  $\mathcal{A}_i$  ( $1 \leq i \leq N$ ) is an ordered node set and can be  $\emptyset$ . Assume  $|\mathcal{A}_i| = L_i$ . Denote the  $k$ th element in  $\mathcal{A}_i$  as  $\mathcal{A}_{i,k}$ . According to the definition of EPA of an ordered node set in Eq. (3), we have

$$\text{EPA}(\mathcal{A}_i) = \sum_{k=1}^{L_i} d_{\mathcal{A}_{i,k}} p_{\mathcal{A}_{i,k}} \cdot \prod_{n=0}^{k-1} (1 - p_{\mathcal{A}_{i,n}}) \quad (34)$$

where  $p_{\mathcal{A}_{i_0}} := 0$ .

For convenience, we handle the ordered set  $\mathcal{A}_i$  as a new node by giving the following definition.

$$p_{\mathcal{A}_i} = 1 - \prod_{n=0}^{L_i} (1 - p_{\mathcal{A}_{i,n}}) \quad \text{and} \quad d_{\mathcal{A}_i} = \frac{\text{EPA}(\mathcal{A}_i)}{p_{\mathcal{A}_i}} \quad (35)$$

If  $\mathcal{A}_i = \emptyset$ , we can choose a positive real number for  $d_{\mathcal{A}_i}$ , s.t.  $d_{b_{i-1}} \geq d_{\mathcal{A}_i} \geq d_{b_i}$ , and  $p_{\mathcal{A}_i} = 0$ .  $d_{b_0}$  is chosen as a positive real number that makes the inequality  $d_{b_0} \geq d_{\mathcal{A}_1} \geq d_{b_1}$  hold.

<sup>3</sup> For simplicity, we denote node using its subscript in this proof.

From Eq. (34) and (35), we can easily derive that

$$d_{\mathcal{A}_{i_1}} \geq d_{\mathcal{A}_i} \geq d_{\mathcal{A}_{i_{L_i}}}, \text{ then } d_{b_{i-1}} \geq d_{\mathcal{A}_i} \geq d_{b_i} \forall 1 \leq i \leq N \quad (36)$$

Now EPA of  $\mathcal{A}$  is represented as

$$\begin{aligned} \text{EPA}(\mathcal{A}) &= \sum_{k=1}^N [d_{\mathcal{A}_k} p_{\mathcal{A}_k} + d_{b_k} p_{b_k} (1 - p_{\mathcal{A}_k})] \\ &\quad \cdot \prod_{n=0}^{k-1} (1 - p_{\mathcal{A}_n}) (1 - p_{b_n}) \end{aligned} \quad (37)$$

where  $p_{\mathcal{A}_0} := 0$  and  $p_{b_0} := 0$ .

Define  $d_{\mathcal{A}_{N+1}} := d_q$  and  $p_{\mathcal{A}_{N+1}} := p_q$ , then

$$\begin{aligned} \Delta A &= \sum_{k=1}^N [p_{\mathcal{A}_k} (d_{\mathcal{A}_k} - d_{\mathcal{A}_{N+1}}) + (1 - p_{\mathcal{A}_k}) p_{b_k} (d_{b_k} - d_{\mathcal{A}_{N+1}})] \\ &\quad \cdot p_{\mathcal{A}_{N+1}} \prod_{n=0}^{k-1} (1 - p_{\mathcal{A}_n}) (1 - p_{b_n}) \end{aligned} \quad (38)$$

$$\Delta B = \sum_{k=1}^N p_{b_k} p_{\mathcal{A}_{N+1}} (d_{b_k} - d_{\mathcal{A}_{N+1}}) \prod_{n=0}^{k-1} (1 - p_{b_n}) \quad (39)$$

So

$$\begin{aligned} \Delta A - \Delta B &= \sum_{k=1}^N [(d_{\mathcal{A}_k} - d_{b_k}) + (1 - p_{b_k}) (d_{b_k} - d_{\mathcal{A}_{k+1}})] \\ &\quad \cdot \left\{ p_{\mathcal{A}_{N+1}} \left[ 1 - \prod_{n=0}^k (1 - p_{\mathcal{A}_n}) \right] \left[ \prod_{n=0}^{k-1} (1 - p_{b_n}) \right] \right\} \geq 0 \end{aligned} \quad (40)$$

As  $\mathcal{A} \supset \mathcal{B}$ ,  $\exists p_{\mathcal{A}_n} \neq 0$  ( $1 \leq n \leq N$ ). We also have  $d_{\mathcal{A}_k} \geq d_{b_k}$  and  $d_{b_k} \geq d_{\mathcal{A}_{k+1}}$ . So inequality (40) holds, that is, inequality (32) holds.

Now we are ready to prove the containing property of  $\mathcal{F}_r^*$ 's. Recall that for arbitrary finite node number,  $N = |\mathcal{C}| \geq 1$ , we want to prove  $\forall \mathcal{F}_{r-1}^*, \exists$  an  $\mathcal{F}_r^*$ , s.t.  $\mathcal{F}_{r-1}^* \subset \mathcal{F}_r^*, \forall 1 \leq r \leq N$ . For each  $r$ , denote one feasible  $\mathcal{F}_r^* = \langle r_1, r_2, \dots, r_r \rangle$  and  $d_{r_1} \geq d_{r_2} \geq \dots \geq d_{r_r}$ . We prove this Lemma by induction.

First, for arbitrary  $N$ , when  $r = 1$ , as  $\mathcal{F}_0^* = \emptyset$ , and  $\mathcal{F}_1^* \neq \emptyset$ , it is obvious that the containing property holds.

Second, when  $r = 2$ , implicating  $N \geq 2$ , we want to prove  $\forall \mathcal{F}_1^*, \exists \mathcal{F}_2^*$ , s.t.  $\mathcal{F}_1^* \subset \mathcal{F}_2^*$ . Assume one feasible  $\mathcal{F}_1^* = \langle 1_1 \rangle$ . We know node  $1_1$  achieves the maximum EPA among the  $N$  nodes. If we assume for any feasible  $\mathcal{F}_2^*$ ,  $\mathcal{F}_1^* \not\subset \mathcal{F}_2^*$ , that is  $1_1 \notin \mathcal{F}_2^*$ , then we can always replace node  $2_2$  as  $1_1$  to achieve at least an equal or larger EPA than  $\text{EPA}(\mathcal{F}_2^*)$ . Then the new node set containing node  $1_1$  should be one feasible node set achieving maximum EPA by selecting 2 nodes from the  $N$  nodes, thus the assumption is wrong, then  $\exists \mathcal{F}_2^*$ , s.t.  $\mathcal{F}_1^* \subset \mathcal{F}_2^*$ . That is, the containing property holds when  $r = 2$ .

Then, we assume  $\forall \mathcal{F}_{m-1}^*, \exists$  an  $\mathcal{F}_m^*$ , s.t.  $\mathcal{F}_{m-1}^* \subset \mathcal{F}_m^*$ , when  $r = m$  ( $m \geq 2$ ), and we want to prove  $\forall \mathcal{F}_m^*, \exists$  an  $\mathcal{F}_{m+1}^*$ , s.t.  $\mathcal{F}_m^* \subset \mathcal{F}_{m+1}^*$ , when  $r = m + 1$ , implicating  $N \geq m + 1$ .

We will prove under the following two cases, the containing property holds.

*Case I:*  $(m + 1)_1 = m_i, i \in \{2, 3, \dots, m\}$ .

Let node  $(m + 1)_j$  ( $2 \leq j \leq m + 1$ ) be the first node in  $\mathcal{F}_{m+1}^*$  but not in  $\mathcal{F}_m^*$ . We can always find such a node because  $\langle (m + 1)_2, \dots, (m + 1)_{m+1} \rangle$  contains more nodes than  $\langle m_i, \dots, m_m \rangle$ . We know that  $\text{EPA}(\mathcal{F}_m^*)$  is the largest EPA achieved by selecting  $m$  nodes, then

$$\text{EPA}(\mathcal{F}_m^*) \geq \text{EPA}(\mathcal{F}_{m+1}^* \setminus \{(m + 1)_j\}) \quad (41)$$

where  $\mathcal{F}_{m+1}^* \setminus \{(m + 1)_j\}$  is the ordered subset obtained by excluding  $(m + 1)_j$  from  $\mathcal{F}_{m+1}^*$ .

From the definition of EPA in Eq. (3), we have

$$\begin{aligned} & \text{EPA}(\langle (m + 1)_j, \mathcal{F}_m^* \rangle) \\ & \geq \text{EPA}(\langle (m + 1)_j, \mathcal{F}_{m+1}^* \setminus \{(m + 1)_j\} \rangle) \end{aligned} \quad (42)$$

Assume  $(m + 1)_{j-1} = m_l$  ( $i \leq l \leq m$ ), we have

$$\begin{aligned} \Delta 1 & := \text{EPA}(\langle m_1, \dots, m_l, (m + 1)_j, m_{l+1}, \dots, m_m \rangle) \\ & \quad - \text{EPA}(\langle (m + 1)_j, \mathcal{F}_m^* \rangle) \\ & = \text{EPA}(\langle m_1, \dots, m_l, (m + 1)_j \rangle) \\ & \quad - \text{EPA}(\langle (m + 1)_j, m_1, \dots, m_l \rangle) \end{aligned} \quad (43)$$

$$\begin{aligned} \Delta 2 & := \text{EPA}(\mathcal{F}_{m+1}^*) \\ & \quad - \text{EPA}(\langle (m + 1)_j, \mathcal{F}_{m+1}^* \setminus \{(m + 1)_j\} \rangle) \\ & = \text{EPA}(\langle (m + 1)_1, \dots, (m + 1)_{j-1}, (m + 1)_j \rangle) \\ & \quad - \text{EPA}(\langle (m + 1)_j, (m + 1)_1, \dots, (m + 1)_{j-1} \rangle) \end{aligned} \quad (44)$$

According to the definition of  $(m + 1)_j$  and the assumption that  $(m + 1)_1 = m_i$ , we have  $\langle (m + 1)_1, \dots, (m + 1)_{j-1} \rangle \subset \langle m_1, \dots, m_l \rangle$ . Then according to inequality (32), we have  $\Delta 1 \geq \Delta 2$ , and combining with inequality (42), we get

$$\text{EPA}(\langle m_1, \dots, m_l, (m + 1)_j, m_{l+1}, \dots, m_m \rangle) \geq \text{EPA}(\mathcal{F}_{m+1}^*) \quad (45)$$

According to the definition of  $\mathcal{F}_{m+1}^*$ , we have

$$\text{EPA}(\langle m_1, \dots, m_l, (m + 1)_j, m_{l+1}, \dots, m_m \rangle) \leq \text{EPA}(\mathcal{F}_{m+1}^*) \quad (46)$$

From inequalities (45) and (46), we have

$$\text{EPA}(\langle m_1, \dots, m_l, (m + 1)_j, m_{l+1}, \dots, m_m \rangle) = \text{EPA}(\mathcal{F}_{m+1}^*) \quad (47)$$

We find the set  $\langle m_1, \dots, m_l, (m + 1)_j, m_{l+1}, \dots, m_m \rangle$  achieving the maximum EPA by using  $m + 1$  nodes, and this set contains  $\mathcal{F}_m^* = \langle m_1, \dots, m_l, m_{l+1}, \dots, m_m \rangle$ .

In the above proof, we implicitly assume  $N \geq m + 1$ . However, it's easy to get that node  $(m + 1)_1$  can not be  $m_i$  when  $N < m + 1$ . Because according to the **Relay priority rule** in Proposition 5.2, when we assume  $(m + 1)_1 = m_i$ , we implicitly exclude nodes  $m_1, \dots, m_{i-1}$  as the forwarding candidates. The number of nodes that can be chosen is  $N - (i - 1) < (m + i) - (i - 1) = m + 1$ , then we do not have enough nodes to choose for  $\mathcal{F}_{m+1}^*$ .

*Case II:*  $(m + 1)_1 \neq m_i, \forall 2 \leq i \leq m$

In this case, we have two sub-cases:

1)  $(m + 1)_1 \neq m_1$

Then  $(m + 1)_1 \notin \mathcal{F}_m^*$ . Assume for any feasible  $\mathcal{F}_m^*, \langle (m + 1)_2, \dots, (m + 1)_{m+1} \rangle \neq \mathcal{F}_m^*$ , we can replace  $\langle (m + 1)_2, \dots, (m + 1)_{m+1} \rangle$  in  $\mathcal{F}_{m+1}^*$  by  $\mathcal{F}_m^*$  to get a sequence  $\langle (m + 1)_1, \mathcal{F}_m^* \rangle$  which achieves at least an equal or larger EPA than  $\text{EPA}(\mathcal{F}_{m+1}^*)$ . So  $\exists \mathcal{F}_{m+1}^*$ , s.t.  $\mathcal{F}_m^* \subset \mathcal{F}_{m+1}^*$ .

2)  $(m + 1)_1 = m_1$

To find  $\langle (m + 1)_2, \dots, (m + 1)_{m+1} \rangle$  is equivalent to selecting  $m$  nodes achieving maximum EPA from the remained node set  $\mathcal{C}_{m_1} \setminus \{m_1\}$  in which all the nodes have advancements larger than  $d_{m_1}$ . As  $\mathcal{F}_m^* \setminus \{m_1\}$  is the node set that achieves the maximum EPA by selecting  $m - 1$  nodes in  $\mathcal{C}_{m_1} \setminus \{m_1\}$ , by the induction assumption, we have  $\mathcal{F}_m^* \setminus \{m_1\} \subset \langle (m + 1)_2, \dots, (m + 1)_{m+1} \rangle$ , then  $\mathcal{F}_m^* \subset \mathcal{F}_{m+1}^*$ .

From the induction above, we know for arbitrary finite node number,  $N \geq 1$ , we have  $\forall \mathcal{F}_{r-1}^*, \exists \mathcal{F}_r^*$  s.t.  $\mathcal{F}_{r-1}^* \subset \mathcal{F}_r^*, \forall 1 \leq r \leq N$ .  $\square$

## References

1. Couto, D., Aguayo, D., Bicket, J., & Morris, R. (2003, September). A high-throughput path metric for multi-hop wireless routing. In *ACM MobiCom'03*, San Diego, California.
2. Zhao, J., & Govindan, R. (2003, November). Understanding packet delivery performance in dense wireless sensor networks. In *ACM Sensys'03*, LA, CA.
3. Zuniga, M., & Krishnamachari, B. (2004). Analyzing the transitional region in low power wireless links. In *IEEE Secon'04*.
4. Zorzi, M., & Rao, R. R. (2003). Geographic random forwarding (geraf) for ad hoc and sensor networks: energy and latency performance. *IEEE Transactions on Mobile Computing*, 2(4), 349–365.
5. Shah, R. C., Bonivento, A., Petrovic, D., Lin, E., van Greunen, J., & Rabaey, J. (2004, November). Joint optimization of a protocol stack for sensor networks. In *IEEE Milcom*.
6. Biswas, S., & Morris, R. (2005, August). Exor: Opportunistic multi-hop routing for wireless networks. In *SIGCOMM'05*, Philadelphia, Pennsylvania.

7. Fussler, H., Widmer, J., Kasemann, M., Mauve, M., & Hartenstein, H. (2003). Contention-based forwarding for mobile ad-hoc networks. *Elsevier's Ad Hoc Networks*, 1(4), 351–369.
8. Zeng, K., Yang, Z., & Lou, W. (2009). Location-aided opportunistic forwarding in multirate and multihop wireless networks. *IEEE Transactions on Vehicular Technology*, 58(6), 3032–3040.
9. Dubois-Ferriere, H., Grossglauser, M., & Vetterli, M. (2007). *Least-cost opportunistic routing*, School of Computer and Communication Sciences, EPFL, Technical Report LCAV-REPORT-2007-001.
10. Raghunathan, V., Schurgers, C., Park, S., & Srivastava, M. (2002). Energy-aware wireless microsensor networks. *Signal Processing Magazine, IEEE*, 19(2), 40–50.
11. Finn, G. G. (1987, March). *Routing and addressing problems in large metropolitan-scale internetworks*, USC/ISI, Technical Report ISI/RR-87-180.
12. Takagi, H., & Kleinrock, L. (1984). Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Transactions on Communications*, 32(3), 246–257.
13. Ko, Y., & Vaidya, N. H. (1998, October). Location aided routing (lar) in mobile ad hoc networks. In *IEEE/ACM MobiCom*.
14. Bose, P., Morin, P., Stojmenovic, I., & Urrutia, J. (1999, August). Routing with guaranteed delivery in ad hoc wireless networks. In *3rd International workshop on discrete algorithms and methods for mobile computing and communications*, Seattle, WA.
15. Karp, B., & Kung, H. (2000, August). Gpsr: Greedy perimeter stateless routing for wireless networks. In *ACM MOBICOM*, Boston.
16. Kuhn, F., Wattenhofer, R., Zhang, Y., & Zollinger, A. (2003, July). Geometric ad-hoc routing: Of theory and practice. In *22nd ACM Symposium on the principles of distributed computing (PODC)*, Boston.
17. Seada, K., Zuniga, M., Helmy, A., & Krishnamachari, B. (2004, November). Energy-efficient forwarding strategies for geographic routing in wireless sensor networks. In *ACM Sensys'04*, Baltimore, MD.
18. Lee, S., Bhattacharjee, B., & Banerjee, S. (2005). Efficient geographic routing in multihop wireless networks. In *MobiHoc*.
19. Zhang, H., & Shen, H. (2010). Energy-efficient beaconless geographic routing in wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 21, 881–896.
20. Bulusu, N., Heidemann, J., & Estrin, D. (2000). Gps-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine*, 7(5), 28–34.
21. Savvides, A., Han, C., & Srivastava, M. B. (2001, July). Dynamic finegrained localization in ad-hoc networks of sensors. In *IEEE/ACM MobiCom*.
22. Zorzi, M., & Amaraoli, A. (2003, October). Advancement optimization in multihop wireless networks. In *Proc. of VTC*.
23. Lin, E., Wolisz, A., & Rabaey, J. (2004, June). Power efficient rendezvous schemes for dense wireless sensor networks. In *IEEE ICC*.
24. El-Hoiydi, A., & Decotignie, J.-D. (2004). Wisemac: An ultra low power mac protocol for the downlink of infrastructure wireless sensor networks. *ISCC*, 2, 244–251.
25. Yang, Z., Zeng, K., & Lou, W. (2009, June). Fsa: A fast coordination scheme for opportunistic routing. In *Communications, 2009. ICC '09. IEEE International Conference on*, pp. 1–5.
26. Shah, R. C., Wietholter, S., Wolisz, A., & Rabaey, J. M. (2005, March). When does opportunistic routing make sense? In *IEEE PerSens*.
27. Shah, R. C., Wietholter, S., Wolisz, A., & Rabaey, J. M. (2005). Modeling and analysis of opportunistic routing in low traffic scenarios. In *WiOpt'05*.
28. Keshavarzin, A., Uysal-Biyikoglu, E., Herrmann, F., & Manjeshwar, A. (2004, March). Energy-efficient link assessment in wireless sensor networks. In *IEEE Infocom*.
29. Shnayder, V., Hempstead, M., Chen, B.-r., Allen, G. W., & Welsh, M. (2004). Simulating the power consumption of large-scale sensor network applications. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, ser. SenSys '04 (pp. 188–200). New York, NY, USA: ACM. [Online] Available: <http://doi.acm.org/10.1145/1031495.1031518>.
30. Ansari, J., Pankin, D., & Mähönen, P. (2009). Radio-triggered wake-ups with addressing capabilities for extremely low power sensor network applications. *IJWIN*, 16(3), 118–130.
31. Zeng, K., Lou, W., Yang, J., & Brown, D. (2007, August). On geographic collaborative forwarding in wireless ad hoc and sensor networks. In *International conference on wireless algorithms, systems and applications (WASA)*, 2007, pp. 11–18.
32. Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to Algorithms* (2nd ed.). Cambridge, Mass: MIT Press.
33. Mica2 mote datasheet, <http://www.xbow.co>.

### Author Biographies



**Kai Zeng** received his Ph.D. degree in Electrical and Computer Engineering at Worcester Polytechnic Institute (WPI) in 2008. He obtained MS in Communication and Information Systems and BS in Communication Engineering both from Huazhong University of Science and Technology, China in 2004 and 2001, respectively. He was a postdoctoral scholar in the Department of Computer Science at University of California, Davis (UCD) from 2008 to

2011. He joined the Department of Computer and Information Science at University of Michigan – Dearborn as an assistant professor in 2011. He is a recipient of the U.S. National Science Foundation Faculty Early Career Development (CAREER) award in 2012. He won Excellence in Postdoctoral Research Award at UCD in 2011 and Sigma Xi Outstanding Ph.D. Dissertation Award at WPI in 2008. His current research interests are in wireless network security, physical layer security, cognitive radio networks, energy efficiency, and cyber-physical systems.



**Jie Yang** received her Ph.D. degree in Electrical and Computer Engineering from Worcester Polytechnic Institute in 2009. She received MS and BS in Communication Engineering from Huazhong University of Science and Technology in 2005 and 2002, respectively. She was a research associate in the Department of Computer Science at University of California, Davis from 2009 to 2010. Her research interests are in wireless networking,

cooperative communication, and energy efficiency.





**Wenjing Lou** is an associate professor at Virginia Polytechnic Institute and State University. Prior to joining Virginia Tech in 2011, she was on the faculty of Worcester Polytechnic Institute from 2003 to 2011. She received her Ph.D. in Electrical and Computer Engineering at the University of Florida in 2003. Her current research interests are in cyber security, with emphases on wireless network security and data security and privacy in cloud computing.

She was a recipient of the U.S. National Science Foundation CAREER award in 2008.