On Geographic Collaborative Forwarding in Wireless Ad Hoc and Sensor Networks*

Kai Zeng   Wenjing Lou   Jie Yang   D. Richard Brown III
ECE Department, Worcester Polytechnic Institute
Worcester, MA 01609
{kzeng, wjlou, abbyyang, drb}@ece.wpi.edu

Abstract

In this paper, we study the geographic collaborative forwarding (GCF) scheme, a variant of opportunistic routing, which exploits the broadcast nature and spatial diversity of the wireless medium to improve the packet delivery efficiency. Our goal is to fully understand the principles, the gains, and the tradeoffs of the node collaboration and its associated cost, thus provide insightful analysis and guidance to the design of more efficient routing/forwarding protocols. We first identify the upper bound of the expected packet advancement (EPA) that GCF can achieve and prove the concavity of the maximum EPA. With energy efficiency as a major concern, we propose a new metric, EPA per unit energy consumption, which balances the packet advancement, reliability and energy consumption. By leveraging the proved properties, we then propose an efficient algorithm which selects a feasible candidate set that maximizes this local metric. We validate our analysis results by simulations, and justify the effectiveness of the new metric by comparing the performance of GCF with those of the existing geographic and opportunistic routing schemes.

1 Introduction

Several recent experimental studies on wireless ad-hoc and sensor networks [3, 14] have shown that wireless links can be highly unreliable and variable. To cope with the randomness of link quality in such networks, a new routing/forwarding scheme, known as opportunistic routing [1, 5, 12, 17, 18], is recently proposed. The general idea behind such scheme is that, for each local transmission, a set of forwarding candidates are selected and one of them is chosen as the actual relay on a per-packet basis depending on its reachability at that instant. Then owing to the spatial diversity of wireless media, the probability of at least one forwarding candidate correctly receiving the packet will increase when more forwarding candidates are involved, which in result will increase the reliability and the expected packet advancement (EPA) to the destination of each transmission.

Although opportunistic routing has shown its effectiveness in achieving better energy efficiency [17,18] and higher throughput [1] than traditional routing. None of the existing works provides a thorough understanding of how well the opportunistic routing can perform and how the selection of the forwarding candidate set will affect the routing efficiency. Questions, such as " a) how many and which neighbor nodes should be involved in the local forwarding? "; and " b) What are the selection criteria and how do they affect the relay priority among the forwarding candidates? ", remain unanswered.

This paper presents the geographic collaborative forwarding (GCF) scheme, which is a variant of opportunistic routing, and studies the candidate selection and prioritizing problems. The main work and contributions of this paper include:

- An upper bound of the EPA that GCF can achieve is derived. We prove that the maximum EPA can only be achieved by giving forwarding candidates closer to the destination higher relay priorities.

- We further prove that the maximum EPA is an increasing and concave function of the number of the forwarding candidates, which indicates that although involving more forwarding candidates will increase the maximum EPA, the EPA gained becomes marginal by doing so.

- With energy efficiency as a major concern, we propose a new metric, *EPA per unit energy consumption*, which strikes a good balance among packet advancement, reliability and energy consumption. Then, by leveraging the proved findings, we propose a localized candidate

---

selection algorithm to maximize the local metric with $O(N^3)$ running time in the worst case and $\Omega(N)$ in the best case, where N is the number of the available next-hop nodes.

We validate our analysis results by simulations, and justify the effectiveness of the new metric by comparing the performance of GCF with those of the existing geographic and opportunistic routing schemes.

The rest of the paper is organized as follows. Section 2 introduces the system model. We generalize the definition of EPA, then identify and prove its principles in Section 3. The trade-off between the EPA and energy consumption in a sensor network scenario is discussed and the new local metric is proposed in Section 4. We propose the candidate selection algorithm in Section 5. Simulation results are presented in Section 6. Section 7 describes the related work, and conclusions are drawn in Section 8.

2 System Model

In this paper, we assume a local GCF scenario like the example in Fig. 1. Assuming node i is forwarding a packet to a sink/destination D, and i_j is one of i's neighbors which is closer to D than i. i is aware of the location[1] of itself, i_j's and D. Let $\mathcal{C}$ denote the set of i_j's which we name as the **available next-hop node set** of node i. Let N=$|\mathcal{C}|$ denote the number of nodes in $\mathcal{C}$. Define d_j in Eq. (1) as the packet advancement to the destination when packet sent by i is relayed by i_j:

$$d_j = Dist(i, D) - Dist(i_j, D) \qquad (1)$$

where $Dist(i, D)$ and $Dist(i_j, D)$ are the Euclidian distance between i and D and between i_j and D, respectively.

Then, each i_j is associated with a pair, (d_j, p_j), where p_j is the *packet reception ratio*[2] (PRR) from node i to i_j. A node is a neighbor of i when PRR from i to it is larger than some non-negligible probability threshold p_T. We assume the PRR on each link is independent. Let $\mathcal{F}$ denote the **forwarding candidate set**, which includes all the nodes involved in the local collaborative 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 opportunistic routing protocols [1,17,18], $\mathcal{F} = \mathcal{C}$. We assume all the nodes in $\mathcal{C}$ and $\mathcal{F}$ are descending ordered[3] according to the advancement s.t. $d_m > d_n, \forall\, m < n$.

The GCF procedure is as following: node i selects $\mathcal{F}$ based on its knowledge of $\mathcal{C}$ (d_j and p_j), and locally broad-

(5, 0.9) (20, 0.4) (30, 0.3) (15, 0.5) (10, 0.85)
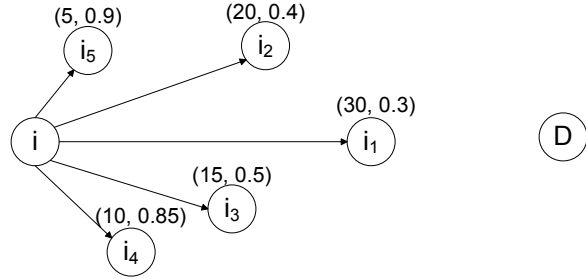
Figure 1. Example in which node i is forwarding a packet to a remote destination D.

casts the packet; when the forwarding candidates in $\mathcal{F}$ receives the packet, they follow a specific priority to relay the packet, that is, 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 the other potential forwarders in $\mathcal{F}$. This procedure iterates until the packet arrives at the destination.

3 EPA and its Principles

Intuitively, involving more forwarding candidates in GCF will bring more chances to the packet to make larger advancement in one transmission. In this section, we will confirm this intuition, and study the relationship between the packet advancement and candidate selection and prioritization. We first generalize the definition of EPA as proposed in geographic routing [9, 11] but applies to arbitrary number of forwarding candidates. We then identify and prove the principles and properties of the EPA, such as **Strictly increasing property**, **Relay priority rule** (which indicates the upper bound of EPA), and **Concavity**, etc.

3.1 EPA Generalization

Let $\pi_j(\mathcal{F}) = \langle i_{j_1}, i_{j_2}, ..., i_{j_r} \rangle$ be one permutation of nodes in $\mathcal{F}$, and the order indicates that nodes will attempt to forward the packet with priority $i_{j_1} > i_{j_2} > ... > i_{j_r}$. We define the EPA for the ordered forwarding candidate set $\pi_j(\mathcal{F})$ in Eq. (2)

$$\text{EPA}(\pi_j(\mathcal{F})) = \sum_{k=1}^{r} d_{j_k} p_{j_k} \cdot \prod_{n=0}^{k-1} \overline{p}_{j_n} \qquad (2)$$

where $\overline{p}_{j_n} = 1 - p_{j_n}$ and $\overline{p}_{j_0}:=1$. The physical meaning of Eq. (2) is the expected packet advancement achieved by GCF in one transmission using the ordered forwarding candidate set $\pi_j(\mathcal{F})$. The EPA metric accurately indicates the relationship between the packet advancement and candidate

---

[1]The node location information can be obtained through localization mechanisms such as [2, 10].

[2]The PRR information can be obtained by using probe messages [3, 7].

[3]In this paper, when we say a node set is ordered, it is descending ordered according to the advancement.

selection and prioritization. Note that when $r = 1$, Eq. (2) degenerates to the "distance×PRR" proposed in geographic routing [9, 11].

3.2 EPA Strictly Increasing Property

Intuitively, increasing the number of forwarding candidates would result in a larger EPA. We present Lemma 3.2 to confirm this intuition.

Definition 3.1 *Define $EM(\mathcal{C}, r)$ be the maximum EPA (defined in Eq. (2)) achieved by selecting r forwarding candidates from $\mathcal{C}$.*

Lemma 3.2 *(Strictly increasing property) $EM(\mathcal{C}, r)$ is a strictly increasing function of r.*

Proof: Assume $1 \leq m < n \leq N$, and without loss of generality, let $\mathcal{A} = \langle i_1, i_2, ..., i_m \rangle$ be the ordered node set achieving $EM(\mathcal{C}, m)$ with forwarding priority $i_1 > ... > i_m$. We then select a subset with $n - m$ nodes from the remaining node set $\{i_{m+1}, i_{m+2}, ..., i_N\}$, say $\mathcal{B} = \langle i_{m+1}, ..., i_n \rangle$. Assume we retain the relay priority of the m nodes in $\mathcal{A}$ and give the nodes in $\mathcal{B}$ lower priorities than those in $\mathcal{A}$. Then in $\mathcal{B}$, we set the nodes with smaller subscripts having higher priorities. So we have

$$EM(\mathcal{C}, n) \geq EPA(\langle i_1, ..., i_n \rangle)$$
$$= EM(\mathcal{C}, m) + EPA(\langle i_{m+1}, ..., i_n \rangle) \prod_{k=1}^{m} \overline{p}_k > EM(\mathcal{C}, m) \quad (3)$$

□

Lemma 3.2 basically indicates that the more nodes get involved in GCF, the larger the EPA can be. The maximum EPA can be obtained by involving all the nodes in $\mathcal{C}$. Then, how to prioritize the candidates to maximize the EPA? We answer this question in the following section.

3.3 Relay Priority Rule

Theorem 3.3 identifies the upper bound of EPA and the corresponding relay priority rule.

Theorem 3.3 *(Relay priority rule) $EM(\mathcal{F}, |\mathcal{F}|)$ defined in Definition 3.1 can only be obtained by giving the node closer to the destination higher relay priority. That is*

$$EM(\mathcal{F}, |\mathcal{F}|) = \sum_{k=1}^{r} d_k p_k \cdot \prod_{n=0}^{k-1} \overline{p}_n \quad (4)$$

where $\overline{p}_0 := 1$.

Proof: We proof Theorem 3.3 by induction on r, the size of $\mathcal{F}$.

First, when r = 1, obviously Eq. (4) holds.

Next, we assume Eq. (4) holds for r = N (N≥1). When $|\mathcal{F}| = N+1$, $\mathcal{F}$ can be divided into $\mathcal{F}_1 = \mathcal{F} - \{i_m\}$ with N nodes and $\mathcal{F}_2 = \{i_m\}$ with 1 node. Then

$$EM(\mathcal{F}, |\mathcal{F}|) = \max_{1 \leq m \leq N+1} \{ \sum_{k=1}^{m-1} d_k p_k \prod_{w=0}^{k-1} \overline{p}_w$$
$$+ \sum_{k=m+1}^{N+1} d_k p_k \frac{\prod_{w=0}^{k-1} \overline{p}_w}{\overline{p}_m} + d_m p_m \frac{\prod_{w=0}^{N+1} \overline{p}_w}{\overline{p}_m} \} \quad (5)$$

Thus we only need to prove for any integer m ($1 \leq m \leq N$),

$$A := \sum_{k=1}^{m-1} d_k p_k \prod_{w=0}^{k-1} \overline{p}_w +$$
$$\sum_{k=m+1}^{N+1} d_k p_k \frac{\prod_{w=0}^{k-1} \overline{p}_n}{\overline{p}_m} + d_m p_m \frac{\prod_{w=0}^{N+1} \overline{p}_w}{\overline{p}_m} \quad (6)$$
$$< B := \sum_{k=1}^{N+1} d_k p_k \prod_{w=0}^{k-1} \overline{p}_w$$

Subtracting A from B, we have

$$B - A = \frac{1}{\overline{p}_m} \sum_{k=m+1}^{N+1} (d_m - d_k) p_m p_k \prod_{w=0}^{k-1} \overline{p}_w > 0 \quad (7)$$

Then the Eq. (4) holds for r = N+1. So it holds for any r ($r \geq 1$). □

Theorem 3.3 indicates that when a forwarding candidate set is chosen, the maximum EPA can only 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. The analysis result is the upper bound of the EPA that GCF can achieve.

Based on the **Relay priority rule**, next, we will identify and prove two important principles about the maximum EPA. 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.

3.4 Containing Property of Feasible Candidate Set

Let $\mathcal{F}_r^*$ be a feasible ordered node set that achieves the $EM(\mathcal{C}, r)$, we have the following containing property of $\mathcal{F}_r^*$'s.

Lemma 3.4 *(Containing property) Given the available next-hop node set $\mathcal{C}$ with N nodes, $\forall \mathcal{F}_{r-1}^*, \exists \mathcal{F}_r^*$, s.t.*

$$\mathcal{F}_{r-1}^* \subset \mathcal{F}_r^* \quad \forall 1 \leq r \leq N \quad (8)$$

Proof: Let $\mathcal{A}=\langle a_1,...,a_M\rangle$[4] be an ordered node set with M nodes, and $\mathcal{B}=\langle b_1,...,b_N\rangle$ with N nodes. $\mathcal{B}\subset\mathcal{A}$ and $b_N=a_M$. For any node $q\notin\mathcal{A}$ with $d_q<d_{a_M}$, we have

$$
\begin{aligned}
&\text{EPA}(\langle\mathcal{A},q\rangle)-\text{EPA}(\langle q,\mathcal{A}\rangle)\\
&>\text{EPA}(\langle\mathcal{B},q\rangle)-\text{EPA}(\langle q,\mathcal{B}\rangle)
\end{aligned}
\tag{9}
$$

We then prove Lemma 3.4 by induction on r.

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.

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 ($m\geq1$). We first prove for any feasible $\mathcal{F}_m^*$ and $\mathcal{F}_{m+1}^*$, the first node in $\mathcal{F}_{m+1}^*$ can not be the nodes from the second place to the last place in $\mathcal{F}_m^*$, that is $(m+1)_1\neq m_i$, $\forall\,2\leq i\leq m$.

We prove this by contradiction. Assume $(m+1)_1=m_i$. Let node $(m+1)_j$ be the first node in $\mathcal{F}_{m+1}^*$ but not in $\mathcal{F}_m^*$. We have

$$
\text{EPA}(\mathcal{F}_m^*)\geq\text{EPA}(\mathcal{F}_{m+1}^*-\{(m+1)_j\})
\tag{10}
$$

then,

$$
\begin{aligned}
&\text{EPA}(\langle(m+1)_j,\mathcal{F}_m^*\rangle)\\
&\geq\text{EPA}(\langle(m+1)_j,\mathcal{F}_{m+1}^*-\{(m+1)_j\}\rangle)
\end{aligned}
\tag{11}
$$

Assume $(m+1)_{j-1}=m_l$, and according to inequality (9), we have

$$
\Delta1>\Delta2
\tag{12}
$$

where

$$
\begin{aligned}
\Delta1 &:=\text{EPA}(\langle m_1,...,m_l,(m+1)_j,m_{l+1},...,m_m\rangle)\\
&\quad-\text{EPA}(\langle(m+1)_j,\mathcal{F}_m^*\rangle)\\
&=\text{EPA}(\langle m_1,...,m_l,(m+1)_j\rangle)\\
&\quad-\text{EPA}(\langle(m+1)_j,m_1,...,m_l\rangle)
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
\Delta2 &:=\text{EPA}(\mathcal{F}_{m+1}^*)\\
&\quad-\text{EPA}(\langle(m+1)_j,\mathcal{F}_{m+1}^*-\{(m+1)_j\}\rangle)\\
&=\text{EPA}(\langle(m+1)_1,...,(m+1)_{j-1},(m+1)_j\rangle)\\
&\quad-\text{EPA}(\langle(m+1)_j,(m+1)_1,...,(m+1)_{j-1}\rangle)
\end{aligned}
\tag{14}
$$

Then combining with inequality (11), we get

$$
\text{EPA}(\langle m_1...m_l,(m+1)_j,m_{l+1}...m_m\rangle)>\text{EPA}(\mathcal{F}_{m+1}^*)
\tag{15}
$$

The inequality (15) contradicts with the fact that $\text{EPA}(\mathcal{F}_{m+1}^*)$ is the largest EPA achieved by selecting m+1 nodes. So the assumption $(m+1)_1=m_i$ is wrong, then $(m+1)_1$ can not be m_i, $\forall\,2\leq i\leq m$. So there are two cases for $(m+1)_1$:

1)$(m+1)_1\neq m_1$. Then $\langle(m+1)_1,\mathcal{F}_m^*\rangle$ should be one $\mathcal{F}_{m+1}^*$.

2)$(m+1)_1=m_1$. By the inductive hypothesis, we have $\mathcal{F}_m^*-\{m_1\}\subset\langle(m+1)_2,...,(m+1)_{m+1}\rangle$, then $\mathcal{F}_m^*\subset\mathcal{F}_{m+1}^*$.

---

[4]For simplicity, we denote node using its subscript in this proof.

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

Lemma 3.4 indicates that an $r-1$-node set that achieves $\text{EM}(\mathcal{C},r-1)$ is a subset of at least one of the feasible r-node sets that achieve $\text{EM}(\mathcal{C},r)$. It also implies that the increasing of the maximum EPA consists with the increasing of the transmission reliability.

3.5 Concavity of Maximum EPA

Following Lemma 3.4, we have the concave property of $\text{EM}(\mathcal{C},r)$ as in Theorem 3.5.

Theorem 3.5 *(Concavity of maximum EPA)*
$EM(\mathcal{C},r+1)-EM(\mathcal{C},r)<EM(\mathcal{C},r)-EM(\mathcal{C},r-1)$, $\forall$ r, s.t. $1\leq r<N$.

Proof: According to Lemma 3.4, assume $\mathcal{F}_{r+1}^*-\mathcal{F}_r^*=\{i_k\}$, and $\mathcal{F}_r^*-\mathcal{F}_{r-1}^*=\{i_j\}$. There are two cases for d_k and d_j.

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

$$
\begin{aligned}
&\mathcal{F}_{r+1}^*=\langle\mathcal{A}_1,i_k,\mathcal{A}_2,i_j,\mathcal{A}_3\rangle,\ \mathcal{F}_r^*=\langle\mathcal{A}_1,\mathcal{A}_2,i_j,\mathcal{A}_3\rangle,\\
&\mathcal{F}_{r-1}^*=\langle\mathcal{A}_1,\mathcal{A}_2,\mathcal{A}_3\rangle
\end{aligned}
\tag{16}
$$

where $\mathcal{A}_i$ ($1\leq i\leq3$) is ordered node set and can be $\emptyset$.

We have

$$
\text{B}:=\text{EPA}(\mathcal{F}_r^*)-\text{EPA}(\langle\mathcal{A}_1,i_k,\mathcal{A}_2,\mathcal{A}_3\rangle)\geq0
\tag{17}
$$

Then,

$$
\begin{aligned}
&[\text{EPA}(\mathcal{F}_r^*)-\text{EPA}(\mathcal{F}_{r-1}^*)]-[\text{EPA}(\mathcal{F}_{r+1}^*)-\text{EPA}(\mathcal{F}_r^*)]\\
&=\text{B}+\overline{p}_{\mathcal{A}_1}\overline{p}_{\mathcal{A}_2}p_kp_j(d_j-\text{EPA}(\mathcal{A}_3))>0
\end{aligned}
\tag{18}
$$

where $\overline{p}_{\mathcal{A}_i}$ is the probability of none of nodes in $\mathcal{A}_i$ receiving the packet correctly.

2) $d_k<d_j$. Similarly, with

$$
\text{B}:=\text{EPA}(\mathcal{F}_r^*)-\text{EPA}(\langle\mathcal{A}_1,\mathcal{A}_2,i_k,\mathcal{A}_3\rangle)\geq0
\tag{19}
$$

we can derive that

$$
\begin{aligned}
&[\text{EPA}(\mathcal{F}_r^*)-\text{EPA}(\mathcal{F}_{r-1}^*)]-[\text{EPA}(\mathcal{F}_{r+1}^*)-\text{EPA}(\mathcal{F}_r^*)]\\
&=\text{B}+\overline{p}_{\mathcal{A}_1}\overline{p}_{\mathcal{A}_2}p_kp_j(d_k-\text{EPA}(\mathcal{A}_3))>0
\end{aligned}
\tag{20}
$$

From the analysis above, we know $\text{EM}(\mathcal{C},r)$ is a concave function of r. $\square$

Combining Lemma 3.2 and Theorem 3.5, we know that giving an available next-hop node set $\mathcal{C}$ 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 many nodes are involved, the gained extra EPA becomes marginal.

4 EPA and Energy Consumption Trade-off

In the following of this paper, we will apply the proved findings about the EPA into a sensor network scenario with energy efficiency as a major concern. First we introduce the energy consumption model. Then we propose a new local metric to balance the EPA and energy consumption.

4.1 Energy Consumption Model

Being energy efficient, we assume nodes only listen to the transmissions intended to themselves. To achieve this, a second low power radio [13] can be used to wake up nodes that should participate in the GCF or to inform the neighbors who (including nodes giving negative advancement) are not selected as forwarding candidates to shut down their data radios. Nodes can also only read the header of packets for early rejection [11]. For simplicity, we only consider the energy consumption of packet transmission and reception. So the total energy consumption for one GCF attempt is:

$$E_t(r) = E_{tx} + r \cdot E_{rx} \qquad (21)$$

where E_{tx} and E_{rx} are the packet transmission and reception energy consumption, respectively.

4.2 EPA per Unit Energy Consumption

Seada and et al [11] define the energy efficiency as the number of packets delivered to the destination per unit of energy consumed by the network, and show that maximizing EPA can achieve optimal energy efficiency in geographic routing. From Lemma 3.2 we know that involving all the available next-hop nodes into forwarding can maximize the EPA in GCF. However, it is not guaranteed to achieve the best energy efficiency. As involving more forwarding candidates also introduces more energy consumption in the network. Clearly there is a trade-off between the per-hop routing efficiency (maximizing EPA) and the overall energy efficiency in GCF.

This trade-off is illustrated in Fig. 2 which is corresponding to the example in Fig. 1 by assuming $E_{tx} = 1$ unit, $E_{rx} = 0.5$ unit. Note that although EM$(\mathcal{C}, r)$ and $E_t(r)$ are both strictly increasing functions of r, the ratio $\frac{\text{EM}(\mathcal{C},r)}{E_c(r)}$ reaches its maximum at $r = 2$, and the corresponding ordered node set is $\langle i_1, i_4 \rangle$ with node i_1 having higher relay priority than i_4.

We then propose a new local metric in Eq. (22), *EPA per unit energy consumption*, which aims to strike a good balance between the routing efficiency and energy efficiency.

$$G(\pi_j(\mathcal{F})) = \frac{\text{EPA}(\pi_j(\mathcal{F}))}{E_t(|\mathcal{F}|)} \qquad (22)$$
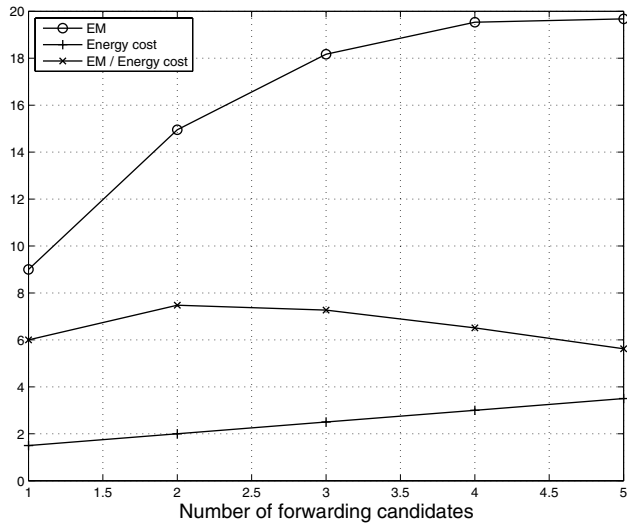
Figure 2. EM, energy cost and their ratio as functions of number of forwarding candidates

The physical meaning of $G(\pi_j(\mathcal{F}))$ is the expected packet advancement to the destination by consuming one unit of energy. The energy cost for successfully delivering one packet from source S to destination D can be approximated by $\frac{Dist(S,D)}{G(\pi_j(\mathcal{F}))}$, then to maximize $G(\pi_j(\mathcal{F}))$ will minimize the energy cost. Now, we formulate the optimization problem as

$$Max \quad G(\pi_j(\mathcal{F})) \quad s.t. \quad \mathcal{F} \subseteq \mathcal{C} \qquad (23)$$

5 Efficient Candidate Selection Algorithm

In this section, we propose an efficient candidate selection algorithm to solve the optimization problem by leveraging the proved principles in Section 3.

5.1 Reformulate the Node Selection Optimization Problem

We know that when r (the size of $\mathcal{F}$) is given, the denominator of the function $G(\pi_j(\mathcal{F}))$ defined in Eq. (22) is fixed, then maximizing $G(\pi_j(\mathcal{F}))$ is equivalent to maximize its numerator. So we can find the suboptimal solution for each $r = 1, ..., N$, then get a global optimal solution by picking the largest one of the suboptimal solutions. Then the optimization problem in (23) is equivalent to

$$Max \quad M(r) := \frac{\text{EM}(\mathcal{C},r)}{E_{tx} + r \cdot E_{rx}} \quad s.t. \quad 1 \le r \le |\mathcal{C}| \qquad (24)$$

We now introduce the following Corollary that can help us solve this optimization problem more efficiently.

```
GetM(C, E_tx, E_rx)
1  M* ← M*_c ← 0; A* ← A*_c ← 0;
2  F* ← F*_c ← ∅; P_{F*} ← P_{F*_c} ← 0; B ← C;
3  while (B ≠ ∅) do /* B is the remained node set */
4    for each node i_j ∈ B
5      F ← F* ⋃{i_j}; P_F ← 1 − P̄_{F*}p̄_j; A ← EPA(F);
6      if A > A*_c || (A = A*_c & P_F > P_{F*_c})
7        A*_c ← A; F*_c ← F; P_{F*_c} ← P_F;
8    end for
9    M*_c ← A*_c/(E_tx + |F*_c| · E_rx);
10   if M*_c ≤ M* /* Local maximum is found */
11     return(M*, F*);
12   else
13     (B, A*, M*, F*, P_{F*})
14     ← (C − F*_c, A*_c, M*_c, F*_c, P_{F*_c});
15 end while
16 return(M*, F*);
```

Table 1. Pseudocode of finding the maximum energy efficiency value M^* and an optimal forwarding candidate set $\mathcal{F}^*$ based on Lemma 3.4 and Corollary 5.1

Corollary 5.1 *(Local maximum of M(r) is global maximum)* If $M(k-1) < M(k)$ and $M(k) \geq M(k+1)$ $(1 \leq k \leq N)$, $M(k) \geq M(k+n)$, $\forall 1 \leq n \leq N-k$.

Proof:

$$M(k) \geq M(k+1)$$
$$\Rightarrow n \times \frac{\text{EM}(\mathcal{C},k+1)-\text{EM}(\mathcal{C},k)}{\text{EM}(\mathcal{C},k)} \leq n \times \frac{E_{rx}}{E_{tx}+k \cdot E_{rx}} \quad (25)$$

Since $\text{EM}(\mathcal{C}, r)$ is concave and positive, we have

$$\frac{\text{EM}(\mathcal{C},k+n)-\text{EM}(\mathcal{C},k)}{\text{EM}(\mathcal{C},k)} \leq n \times \frac{\text{EM}(\mathcal{C},k+1)-\text{EM}(\mathcal{C},k)}{\text{EM}(\mathcal{C},k)} \quad (26)$$

From inequality (25) and (26), we can derive

$$\frac{\text{EM}(\mathcal{C},k)}{E_{tx}+k \cdot E_{rx}} \geq \frac{\text{EM}(\mathcal{C},k+n)}{E_{tx}+(k+n)E_{rx}} \quad (27)$$

that is $M(k) \geq M(k+n)$, $\forall 1 \leq n \leq N-k$. $\square$

5.2 Efficient Node Selection Algorithm

Based on the **Containing property** in Lemma 3.4, a straightforward way to find an optimal node set containing r nodes is to add a new node into the optimal node set containing r-1 nodes. Furthermore, when a local maximum is found, it is the global maximum according to Corollary 5.1. The algorithm GetM in Table 1 finds an optimal forwarding candidate set $\mathcal{F}^*$ and the corresponding energy efficiency value M^* of the objective function defined in (24). Note that $\mathcal{F}^*$, $\mathcal{F}^*_c$ and $\mathcal{F}$ are all ordered node sets with nodes

closer to the destination having higher relay priorities. For feasible sets having the same maximum EPA, we choose the one that achieves higher one-hop transmission reliability (line 6).

It's not difficult to find an algorithm to calculate EPA($\mathcal{F}$) (in line 5) in $O(|\mathcal{F}|)$ running time. Then the algorithm GetM costs $O(N^3)$ running time in the worst case, and $\Omega(N)$ in the best case.

6 Performance Evaluation

In this section, we evaluate the performance of GCF through simulations. We compare GCF with the geographic routing which only has one forwarding candidate that achieves the maximum EPA, and the opportunistic routing which involves all the available next-hop nodes as forwarding candidates.

6.1 Evaluation Metrics

We define the following evaluation metrics:

- *Packet Delivery Ratio (PDR)*: percentage of packets sent by the source that actually reach the sink. This is a measure for reliability.

- *Hop count*: it is measured as the number of hops a successfully delivered packet travels from source to destination.

- *Energy efficiency* $\eta(S, D)$: this metric is measured in bit-meters per Joule (bmpJ). It is calculated as in Eq. (28),

$$\eta(S, D) = \frac{L_{pkt} \cdot N_s \cdot PDR \cdot Dist(S, D)}{E_{total}} \quad (28)$$

where N_s denotes the number of packets sent out from source S and E_{total} is the (transmission and reception) energy consumed by all the nodes involved in the routing procedure excluding the sink.

6.2 Simulation Setup

To simulate a realistic channel model for lossy WSNs, we use the log-normal shadowing path loss model derived in [19]. Non-coherent FSK and Manchester are used as the modulation and encoding schemes (ρ=2), respectively. The path-loss exponent α = 3.5, log-normal shadowing variance σ = 4, and p_T (in Section 2) is set to 0.1. The transmission and reception energy consumption are as following.

$$E_{tx} = T_t \cdot (\frac{P_{trans}}{\gamma} + P_{elec}), \ \ E_{rx} = T_t \cdot P_{elec} \quad (29)$$

(a) PDR vs node density   (b) Hop count vs node density   (c) Energy efficiency vs node density
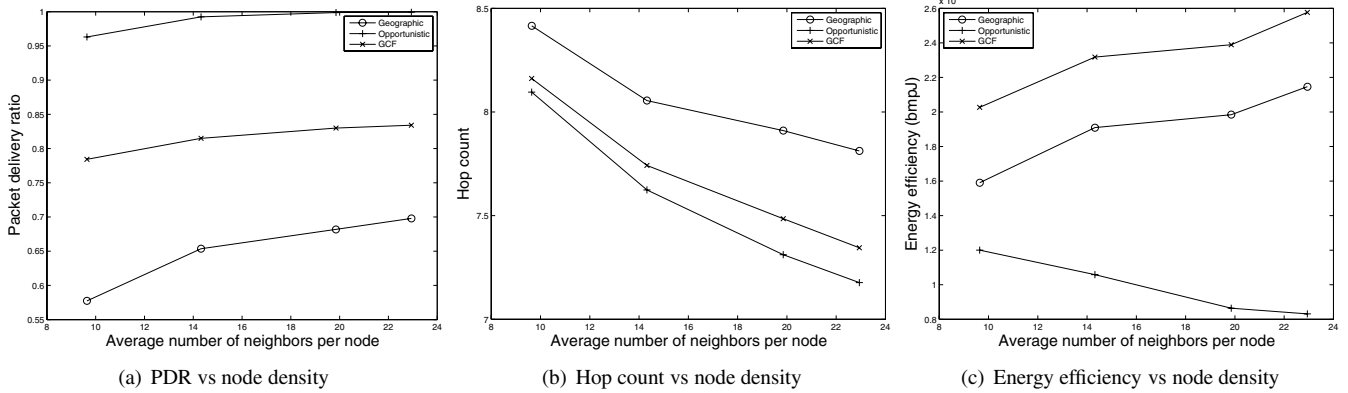
Figure 3. Simulation results under different node densities

where T_t is the packet transmission time, P_{elec} is the power of electronic, and γ is the power amplifier efficiency which is set to be 0.3 in our simulation.

The simulated sensor network has stationary nodes uniformly distributed in a 60×60 m^2 square region, with nodes having identical fixed transmission power of 0 dbm. The frame length is fixed on 50 bytes with preamble of 20 bytes. The source and the sink node are fixed at two corners across the diagonal of the square area. All simulations are run for 5000 iterations. For each iteration, node locations are randomly re-assigned and PRRs between nodes are re-calculated. There is no retransmission applied in the simulations. We vary the average number of neighbors per node as $9.5, 14, 20, 23$ to investigate the impact of node density on the performance of these three protocols. The reception power consumption is fixed on 2mw, so the reception to transmission energy ratio is $\frac{3}{8}$.

6.3 Simulation Results

From Fig. 3(a), it's not surprising to see that the opportunistic routing has the highest PDR, the geographic routing has the lowest, and GCF is in the middle, as they involve all, one and some of the neighbor nodes into the local forwarding, respectively. It's the same reason for the hop count results in Fig. 3(b). Fig. 3(c) shows that GCF achieves the best energy efficiency among the three schemes. Another interesting observation is that the energy efficiency of GCF and the geographic routing is increased as the network becomes denser, while the opportunistic routing shows the opposite trend. The reason is that although the hop count of the opportunistic routing is decreased when node density increases, the energy consumption due to unnecessarily involving more nodes in forwarding overwhelms the benefit of hop count decreasing.

7 Related Work

7.1 Geographic Routing

Owing to its scalability, statelessness, and low maintenance overhead, geographical routing is considered as an efficient paradigm for data forwarding in multi-hop wireless ad hoc and sensor networks. Early works [4, 6, 8] on geographic routing exploit the concept of maximum advancements towards the destination to route packets in a greedy manner. However, recent empirical measurements [3, 14] have proved that the unit disk connectivity model, on which these solutions are based, often fails in real settings. More recent works on geographic routing are focused on lossy channel situations. Seada, et al. [11] articulated the distance–hop energy trade-off for geographic routing. They concluded that packet advancement timing packet reception ratio, the EPA, 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 [16]. Lee et al. [9] 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, NADV only applies to geographic routing which involves single forwarding candidate and can not be directly used for GCF.

7.2 Opportunistic Routing

Opportunistic routing takes advantage of the broadcast nature and spatial diversity of the wireless medium to forward packets more efficiently than traditional routing protocols. Some variants of opportunistic routing, such as ExOR [1] and opportunistic any-path forwarding [15], rely on the

global knowledge of the network to select candidates and prioritize them. While some other variants [5,12,17,18] use the location information of nodes to define the candidate set and relay priority. In general, a node's location will affect its chance of being selected as the actual relay in this kind of opportunistic routing, however, the particular choice of the relay node depends on the policies of a specific protocol variant. In GeRaF [17, 18], the next-hop neighbors of the current forwarding node are divided into sets of priority regions with nodes closer to the destination having higher relay priorities. Similar to [17,18], in [12], 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. [5] discussed three suppression strategies of contention-based forwarding to avoid packet duplication in mobile ad hoc networks.

None of these works provides good answers to the questions such as "how well can opportunistic routing perform?" and "how do the selection and prioritization of the forwarding candidates affect the routing efficiency?". Our work answers these questions and provides more insightful understanding of the trade-off among the packet advancement, energy cost and reliability associated with the node collaboration.

8   Conclusion

In this paper, we study the principles and properties of GCF scheme. We indicate the upper bound of the EPA that GCF can achieve and the corresponding relay priority rule, that is, the maximum EPA can only be achieved by giving the forwarding candidates closer to the destination higher relay priorities. We also show that the maximum EPA achieved by selecting r nodes is a strictly increasing and concave function of r. With the energy efficiency as the major concern, we propose a new local metric, *EPA per unit energy consumption*, which balances the packet advancement, reliability and energy consumption in a good way. Then, by leveraging the proved findings, we propose an efficient localized candidate selection algorithm which determines the forwarding candidate set that maximizes the proposed new metric. The simulation results validate our analysis and justify that GCF achieves better energy efficiency than the corresponding geographic and opportunistic routing schemes.

References

[1] S. Biswas and R. Morris. Exor: Opportunistic multi-hop routing for wireless networks. In *SIGCOMM'05*, Philadelphia, Pennsylvania, Aug. 2005.

[2] N. Bulusu, J. Heidemann, and D. Estrin. Gps-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine*, 7(5):28–34, Oct. 2000.

[3] D. Couto, D. Aguayo, J. Bicket, and R. Morris. A high-throughput path metic for multi-hop wireless routing. In *ACM MobiCom'03*, San Diego, California, Sept. 2003.

[4] G. G. Finn. Routing and addressing problems in large metropolitan-scale internetworks. Technical Report ISI/RR-87-180, USC/ISI, March 1987.

[5] H. Fussler, J. Widmer, M. Kasemann, M. Mauve, and H. Hartenstein. Contention-based forwarding for mobile ad-hoc networks. *Elsevier's Ad Hoc Networks*, 1(4):351–369, Nov. 2003.

[6] B. Karp and H. Kung. Gpsr: Greedy perimeter stateless routing for wireless networks. In *ACM MOBICOM*, Boston, August 2000.

[7] A. Keshavarzin, E. Uysal-Biyikoglu, F. Herrmann, and A. Manjeshwar. Energy-efficient link assessment in wireless sensor networks. In *IEEE Infocom*, March 2004.

[8] F. Kuhn, R. Wattenhofer, Y. Zhang, and A. Zollinger. Geometric ad-hoc routing: Of theory and practice. In *22nd ACM Symposium on the Principles of Distributed Computing (PODC)*, Boston, July 2003.

[9] S. Lee, B. Bhattacharjee, and S. Banerjee. Efficient geographic routing in multihop wireless networks. In *MobiHoc*, 2005.

[10] A. Savvides, C. Han, and M. B. Strivastava. Dynamic finegrained localization in ad-hoc networks of sensors. In *IEEE/ACM MobiCom*, July 2001.

[11] K. Seada, M. Zuniga, A. Helmy, and B. Krishnamachari. Energy efficient forwarding strategies for geographic routing in wireless sensor networks. In *ACM Sensys'04*, Baltimore, MD, Nov. 2004.

[12] R. C. Shah, A. Bonivento, D. Petrovic, E. Lin, J. van Greunen, and J. Rabaey. Joint optimization of a protocol stack for sensor networks. In *IEEE Milcom*, Nov. 2004.

[13] N. H. Vaidya and M. J. Miller. A mac protocol to reduce sensor network energy consumption using a wakeup radio. *IEEE Transactions on Mobile Computing*, 4(3):228–242, 2005.

[14] J. Zhao and R. Govindan. Understanding packet delivery performance in dense wireless sensor networks. In *ACM Sensys'03*, LA,CA, Nov. 2003.

[15] Z. Zhong, J. Wang, and S. Nelakuditi. Opportunistic anypath forwarding in multi-hop wireless mesh networks. Technical Report TR-2006-015, USC-CSE, 2006.

[16] M. Zorzi and A. Armaroli. Advancement optimization in multihop wireless networks. *Proc. of VTC*, Oct. 2003.

[17] M. Zorzi and R. R. Rao. Geographic random forwarding (geraf) for ad hoc and sensor networks: energy and latency performance. *IEEE Transactions on Mobile Computing*, 2(4), 2003.

[18] M. Zorzi and R. R. Rao. Geographic random forwarding (geraf) for ad hoc and sensor networks: multihop performance. *IEEE Transactions on Mobile Computing*, 2(4), 2003.

[19] M. Zuniga and B. Krishnamachari. Analyzing the transitional region in low power wireless links. In *IEEE Secon'04*, 2004.