SPECIAL ISSUE PAPER

Energy-efficient beaconless geographic routing in energy harvested wireless sensor networks

Oswald Jumira1,*,†, Riaan Wolhuter1 and Sherali Zeadally2

1Department of Electrical and Electronic Engineering, University of Stellenbosch, Stellenbosch, South Africa
2Department of Computer Science and Information Technology, University of the District of Columbia, Washington DC 20008, USA

SUMMARY

We propose a routing scheme called energy-efficient beaconless geographic routing with energy supply (EBGRES) for wireless sensor networks. EBGRES provides loop-free, fully stateless, energy-efficient source-to-sink routing with minimal communication overhead without the help of prior neighborhood knowledge. It locally determines the duty-cycle of each node, based on an estimated energy budget for each period, which includes the currently available energy, the predicted energy consumption and the energy expected from the harvesting device. In EBGRES, each node sends out the data packet first rather than a control message. By sending a data packet first, EBGRES performs the neighbor selection only among those neighbors that successfully received the data packet. EBGRES uses a three-way (DATA/ACK/SELECT) handshake and a timer-assignment function, the Discrete Dynamic Forwarding Delay (DDFD). We investigate the lower and upper bounds on hop count and the upper bound on energy consumption under EBGRES for source-to-sink routing. We further demonstrate the expected total energy consumption along a route toward the sink with the proposed EBGRES approach including a lower bound on energy consumption when the node density increases. Copyright © 2012 John Wiley & Sons, Ltd.

Received 13 January 2012; Revised 21 February 2012; Accepted 25 February 2012

KEY WORDS: communication; energy; sensor networks; routing; wireless

1. INTRODUCTION

Research on data routing strategies for wireless sensor networks (WSNs) has focused on energy-efficiency [1–6] with interest increasingly centered on real-time applications, which involve time-critical data and increased network lifetime [7]. The dramatic increase in WSN applications that employ many different kinds of sensors has made it necessary to develop routing protocols that can handle a wide range of data types efficiently, maintain connectivity over long periods of time, and achieve timely delivery of sensed data. Most proposed energy-efficient routing schemes continue to be constrained by the energy supply, which in most cases has been battery based. This has a negative impact on the lifetime of the wireless sensor network and network reliability for different applications. WSNs require new types of power sources and low-latency routing schemes that can accommodate these challenges. In applications such as real-time monitoring for which a long network lifetime and high quality services are required, the batteries that power the nodes need to be replaced or recharged, but because of environmental constraints this is not possible in all cases. For most wireless sensor nodes batteries are the most prevalent powering method because they are

*Correspondence to: Oswald Jumira, MIH Media Lab, Department of Electrical and Electronic Engineering, University of Stellenbosch, Stellenbosch, South Africa.
†E-mail: oswald@ml.sun.ac.za

Copyright © 2012 John Wiley & Sons, Ltd.
a cost effective, ubiquitous, commonly known, and well understood powering technology. However, they present specific challenges that include finite useful life, replacement cost, and disposal concerns. Although they are an ideal solution for many applications, there are many other applications where batteries fail to fit application requirements: for example, the asset is not available to replace the batteries, the cost of battery replacement is too expensive over the life of the product, the device is in a hazardous environment, or the device is embedded and a continuous power supply is required. Applications with these needs provide a good fit for receiving power via ambient energy harvesting [8].

Unlike the microprocessor industry or the communication hardware industry, where the computation capability or the line rate has been continuously improved (almost doubled every 18 months), battery technology has been relatively unchanged for many years [8]. Ambient energy harvesting as a power solution has steadily gained momentum in recent years, especially with significant progress in the functionality of low power embedded electronics such as wireless sensor nodes. We define an energy harvesting node as any system that draws part or all of its energy from the environment such as solar energy, temperature variations, kinetic energy or vibrations. A key distinction of this energy from that stored in the battery is that this energy is potentially infinite, although there may be a limit on the rate at which it can be used. Energy harvesting sensor nodes have either an onboard energy harvesting component as shown in Figure 1, or a sensor node can be connected to an energy harvesting device/component to form one device. In our research we look at solar-based energy harvesting, which has a diurnal characteristic (available readily during the day and no sunlight at night). By generating power from environmental energy, the dependency on batteries can be reduced or even eliminated [7, 9, 10]. Commercially available energy harvesting devices have been developed by Texas Instruments [11] and Microstrain [12]; however, they are currently expensive but the prices are predicted to fall dramatically in the next 5 years [8].

However, limited attention has been given to routing within a network of sensor nodes running on environmental energy. Once a sensor has been deployed, it must be able to operate as autonomously as possible. Each sensor node in an energy-harvesting WSN must be aware that the amount of environmental energy it can gather depends on the time, location, and the surroundings it operates in. This awareness of surroundings should influence the way it operates, for example, operation of a solar harvesting node will adapt to changes in surrounding weather. Each sensor node must determine the appropriate duty cycle (DC) with which a sensor node can operate perpetually [7]. A duty cycle is the time that a sensor node spends in an active state as a fraction of the total time under consideration to manage energy effectively. Knowledge of energy-harvesting devices characteristics should be incorporated in its routing scheme. Research has been carried out in this field over the years [4, 13–16]. The energy-aware routing schemes that have been proposed in recent years are not localized, not scalable, rely on the dissemination of route discovery information and routing tables, have limited lifetime, and they rely on beacons for dynamic networks changes. To address these limitations, we need to investigate and propose energy-aware routing schemes that result in increased energy-efficiency, reliability, scalability, and network lifetime.

Geographic routing, in which each node forwards packets based only on its location, its direct neighbors, and the destination, is particularly attractive to resource-constrained sensor networks.

**Figure 1.** Key components of an energy harvesting wireless sensor node.
Centralized solutions generally need global knowledge, including the position and activity status of all network node; thus, nodes need the dissemination of route discovery information and need to maintain routing tables. To avoid this communication overhead, we focus on localized algorithms where only local information such as the position of the current node holding a packet, the position of its 1-hop neighbors, and that of the destination are required. The localized nature of geographic routing eliminates the overhead caused by route establishment and maintenance. By minimizing such overheads, memory usage is kept to a minimum, which also increases scalability.

In conventional geographic routing schemes nodes are required to maintain more or less accurate position information of all its direct neighbors, and the position of a node is made available to its direct neighbors by periodically broadcasting beacons [14, 17]. If WSNs network topology does not change much, maintaining neighbor information can greatly improve the performance because of the reusability of the stored information and the low maintenance cost. In many other application scenarios (e.g., animal tracking and monitoring), the network topology may frequently change because of node mobility, node status (duty cycle), node or link faults, etc. For these application scenarios routing protocols based on maintaining neighbor information suffer from several drawbacks. First, the maintenance of neighbor information incurs too much communication overhead and results in significant energy expenditure. Second, the collected neighbor information can quickly get outdated, which, in turn, leads to frequent packet drops. Third, the storage and maintenance of neighbor information consumes memory, which is also a scarce resource in WSNs.

To overcome the challenges of conventional geographic routing schemes in scenarios with dynamic network changes, beaconless geographic routing protocols [17–22] have been proposed. Beaconless routing schemes, in which each node forwards packets without the help of beacons and without the maintenance of neighbor information, are fully reactive. Reactive protocols or routing schemes establish a communication path only when a node has information to send; the routing functions are executed only at that moment Beacons can interfere with regular data transmission (causing collisions and congestions), which leads to bandwidth and energy wastage, especially for those nodes not taking part in the routing task [22]. When a node has a packet to transmit, it broadcasts the packet to its neighbors. The most suitable neighbor to relay the packet is determined based on the contention mechanism in which each neighbor determines a proper delay to forward the packet further based on how well it is suited as the next-hop relay. Therefore, beaconless routing schemes are robust to topology changes because the forwarding decision is based on the actual topology at the time a packet is to be forwarded. In most existing beaconless routing schemes such as contention based forwarding [18], beaconless routing (BLR) [19] and energy-efficient beaconless geographic routing (EBGR) [20], each node forwards packets based on hop-count routing metrics (e.g., each node selects its neighbor closest to the destination as its next-hop relay). These routing metrics are simple to implement, but they cannot guarantee energy efficiency, which is a major concern in WSNs. Most beaconless protocols in the literature assume a perfect wireless channel [18–20] and do not consider the problems of losses and interferences created by transmission of messages during the next forwarding process.

1.1. Contributions of this work

In this paper, we present a routing protocol called energy-efficient beaconless geographic routing with energy supply (EBGRES). EBGRES addresses the problem of providing energy-efficient beaconless geographic routing with energy supply for dynamic wireless sensor networks in which network topology frequently changes over time. Without any prior knowledge of neighbors, EBGRES aims to minimize the total energy consumed while delivering each packet to the sink. Unlike previous schemes, EBGRES starts sending out the data packet itself rather than a control message. The basic rationale for doing so is that clear-to-send and return-to-send messages can traverse a link that data packets cannot. In EBGRES energy-harvesting at the nodes offers more residual energy for the transmission of data packets. By sending a data packet first (DATA), EBGRES performs the neighbor selection using greedy forwarding only among those neighbors that successfully received the data packet that fall in the relay search region. EBGRES uses a three-way (DATA/ACK/SELECT) handshake and a timer-assignment contention mechanism, the discrete dynamic forwarding delay
(DDFD), which was proposed in [22], for selecting and confirming the next forwarding node within the relay search region. DDFD divides the neighbor area into subareas according to the progress of the packet towards the destination and this enables the delay timer for receiving the ACK messages at the source. The nodes that are selected for forwarding the data will also make forwarding decisions based on the available energy (from energy harvesting) and duty cycle updates to increase the network lifetime and workload. If there is no node in the relay search region, the forwarding node enters into beaconless recovery mode (perimeter routing is when a packet reaches a region where greedy forwarding is impossible, the algorithm recovers by routing around the perimeter of the region [23,24]) from a local minimum, which are the boundary void areas. The key contributions of this paper are as follows:

- We propose an online geographic routing scheme called EBGRES, which can provide fully stateless, energy-efficient source-to-sink and scalable routing with low communication overheads without maintaining neighborhood information with perpetual energy supply from a solar based energy harvester.
- We prove that EBGRES is loop-free in a greedy forwarding mode, and we establish the lower and upper bounds on hop count for source-to-sink routing for WSNs.
- We establish the upper bound on energy consumption for source-to-sink data delivery under EBGRES, assuming packet loss and link failures in greedy forwarding mode as the number of nodes increases.

The rest of the paper is structured as follows: The related works on energy harvesting, energy-efficient routing and beaconless geographic routing are discussed in Section 2. The system network, energy harvesting and energy consumption models are described in Section 3. We describe EBGRES in Section 4. In Section 5 we present a detailed theoretical analysis for EBGRES. Finally Section 6 makes some concluding remarks.

2. RELATED WORKS

There has been significant research in alternative methods of providing external consistent energy to wireless sensor nodes by utilizing energy harvesting technologies. Several harvesting techniques have been demonstrated [4, 15, 25] many of which provide the ability to extract energy from the environment. These techniques must be efficiently integrated into the design and operation of the underlying wireless sensor network. Therefore, it is imperative to understand the sensor application requirements and design a suitable energy harvesting product that fits their application with minimal redesign and development. The improvement in energy-efficiency through routing schemes and increased energy supply from the environment through energy harvesting achieves greater network lifetime, increased data sampling, increased workload and opens up other application areas for sensor networks. Energy-aware methods that make routing decisions have been considered previously [5,15,16,26]. Various methods have been proposed to collect the residual battery status of a distributed system [14] and also to estimate the future energy consumption at various nodes [25].

There has been a lot of research on geographic routing because it has been demonstrated to be energy efficient [5, 24, 25, 27]. The Most Forwarding progress within Radius protocol which was proposed in [28] is one of the earliest geographic routing algorithms in which each node forwards its packets to the neighbor that has the maximum progress to the destination. However the challenge was that the MRF scheme cannot guarantee that packets are delivered in an energy-efficient manner. Over the last few years a lot of research has been undertaken to improve energy efficiency in geographic routing. In [29] the Geographical Power Efficient Routing (GPER) protocol for sensor networks was presented. Each sensor node makes local decisions as to how far to transmit; therefore, the protocol is power efficient, highly distributed, and scalable. GPER does not guarantee increased network lifetime, support increasing workload, and network reliability. A novel geographic routing protocol called Energy-Efficient Geographic Routing (EEGR) for WSNs was proposed in [17], which can provide near-optimal energy-efficient routing based only on local information. In EEGR both the geographic information and the transceiver power characteristics are employed to make forwarding decisions, thereby enabling an energy-aware localized routing strategy. EEGR makes
use of beacons, does not take into account realistic channel conditions, and does not guarantee increased network lifetime and increasing workload. In [23], geographic and energy aware routing is proposed, which is a geographic routing protocol that takes into account a nodes’ residual energy information. Geographic and energy aware routing uses energy-aware and geography-based neighbor heuristics to route a packet towards the target region but it does not take into account realistic wireless channel conditions and environmental energy supply.

An energy-efficient geographic routing protocol in environmentally powered WSNs is studied in [14]. The authors found that the protocol maintains a higher minimum residual energy on nodes and achieves better load balancing in terms of having a smaller standard deviation of residual energy among nodes. The protocol does not consider the property of the energy changing (recharging and consumption). In [3], an opportunistic routing protocol, Energy Harvesting Opportunistic Routing (EHOR) is designed for routing in multihop WSNs powered solely using ambient energy harvesters (WSN-HEAP). Opportunistic routing allows any node (in the forwarder list) that overhears the transmission and is closer to the destination to participate in the forwarding of the packet. EHOR takes into account energy constraints because nodes have to shut down to recharge once their energy is depleted. EHOR increases goodput and efficiency compared with traditional opportunistic routing protocols and other nonopportunistic routing protocols suited for WSN-HEAP. EHOR does not take into account a realistic channel and is highly unreliable. The main challenge for EHOR is how to select the appropriate forwarder list such that the expected energy cost is minimized. In [16], a new adaptive algorithm, Distributed Energy Harvesting Aware Routing Algorithm (DEHAR), for finding energy optimised routes in a wireless sensor network with energy harvesting is proposed. The algorithm finds an energy efficient route from each source node to a single sink node, taking into account the current energy status of the network. DEHAR has high overheads because it is a centralized solution that generally needs global knowledge, including position and activity status of all network nodes. As a result nodes need the dissemination of route discovery information and need to maintain routing tables.

The above energy-efficient geographic routing schemes utilize the beaconing mechanism which has some issues, such as generating interferences with regular data transmissions and consuming bandwidth and battery power. The algorithms are mostly based on unrealistic wireless channels and assume successful greedy forwarding. The lifetime of the network is limited to the battery life and the energy management. The sensor nodes are generally based on event driven applications and do not perform well under realtime and multimedia applications because of sustainable energy challenges, increased overheads and network reliability. Our proposed algorithm EBGRES address these challenges leading to increased network performance and network lifetime because of perpetual energy supply.

Beaconless routing has been proposed to deal with a dynamic network topology in which a node forwards packets without maintaining neighbor information. The first contention-based georouting algorithms were proposed under the names BLR [19] and contention-based forwarding [18]. They all follow the same basic principle: first the forwarder broadcasts the control message to its (unknown) neighbors. Some of the neighbors, which are located in the forwarding area, are candidates for the next hop and contend for the message (contention process). They set a timer in accordance to their distance to the destination. The timer is determined by a delay function depending on the progress made to the destination. Once the first timer expires, the respective candidate retransmits the message. The other candidates overhear the retransmissions and cancel their scheduled transmissions (suppression process). In [22] Sanchez et al. proposed the Beaconless on Demand Strategy for Geographic Routing in wireless sensor networks (BOSS). BOSS uses a three-way handshake similar to IEEE802.11 (return-to-send/clear-to-send) handshake and timer-assignment function, which divides the neighbor area into subareas according to the progress towards the destination and helps in reducing collisions. However, most of the proposed beaconless schemes reported in the literature employ hop-count-based routing metric, which is not efficient in terms of energy consumption. A beaconless routing algorithm that uses a cost-over progress approach to determine energy-optimal links was proposed in [30]. To address the problem of providing energy-efficient beaconless geographic routing for dynamic wireless sensor networks in which the network topology frequently
changes over time, a novel routing protocol called energy-efficient beaconless geographic routing (EBGR) is designed [20]. EBGR assumes an unrealistic channel and no failure in greedy forwarding. The energy supply is finite leading to limited lifetime and workload compared with EBGRES. The performance is not adaptable to an infinite energy supply.

Various geographic routing protocols use greedy forwarding as the basic mode of operation. Greedy forwarding may fail when a node cannot find a better neighbor than itself to forward the packet. To recover from a local minimum, which is the boundary void areas where greedy forwarding fails, greedy-face-greedy (GFG) [17], greedy perimeter stateless routing [31], and GOAFR+ [32] route a packet around the faces of a planar subgraph (e.g., relative neighborhood graph (RNG) and Gabriel graph (GG)) when a local minimum is encountered. Planar graphs are constructed based on neighborhood information, which is not prior knowledge in beaconless routing schemes. In BLR [19], a Request-response approach was proposed to determine the local minima. In [24], algorithms for constructing different proximity graphs in beaconless routing were designed. For guaranteed delivery in WSNs, most existing geographic routing algorithms [19, 23, 24] switch between the greedy forwarding mode and recovery mode depending on the network topology.

Energy-efficient beaconless geographic routing with energy supply strives to guarantee packet delivery and reduces the energy consumption of both greedy and recovery parts with an increase in energy supply (from energy harvesting) to increase lifetime, network reliability, and support increased workload. EBGRES possesses the following distinguishing characteristics:

- **Localized**: To make an end-to-end routing decision, a node has to be aware only of its location, of its neighbors, and of the final destination.
- **Scalable**: The end-to-end path is memoryless because no routing information needs to be stored at the forwarding node leading to a scalable protocol where no information is embedded in the message.
- **Loop free**: The end-to-end path is loop-free because the greedy step always chooses among its neighbors in the forwarding direction of the destination (nodes closer to the destination than itself). This makes any sender nodes on the path forward to node closer to the destination than the sender node.
- **Guarantees delivery**: The end-to-end path uses two routing phases. The first phase uses a localized greedy protocol that is resistant to routing failure. End-to-end guaranteed delivery is ensured by invoking the second routing phase — that is, perimeter routing is invoked when needed.
- **Energy efficient**: Every routing step along the end-to-end path takes an energy aware route based on energy harvesting rate of the sensor nodes. To avoid expensive long edges, the end-to-end path computes an energy weighted localized shortest path from the relaying node to all its neighbors in the forward direction and selects the one that minimizes the cost of the shortest path to the progress towards the destination.

3. MODELS

3.1. Network model

Given a set of nodes distributed in space, we need to specify which nodes can receive the transmission of a node. Without loss of generality, it is assumed that no two nodes are located at the same position. Throughout this paper, if a node \( u \) is within a node \( v \)'s transmission range we say that \( u \) is adjacent to \( v \), or equivalently, that \( u \) is a neighbor of \( v \). In the absence of interference, this relationship is typically symmetric (or undirected); that is, if a node \( u \) can hear a node \( v \), \( v \) can also hear \( u \). All nodes are equipped with the same radio transceiver that enables a maximum transmission range \( R \). Each node knows its own location and the location of the sink. The connectivity model we will use is the unit disk graph [33]. Nodes having omnidirectional radio antennas are assumed to be deployed in a planar, unobstructed environment. In this model, any two nodes \( u \) and \( v \) can communicate with each other reliably if and only if \( |uv| \leq R \), where \( |uv| \) is the Euclidean distance between \( u \) and \( v \). On the basis of a realistic communication model in which data loss is estimated
by packets reception rate, we extend our scheme to achieve localized energy-efficient beaconless routing in the presence of unreliable communication links and environmental energy supply.

3.2. Energy harvesting and duty cycle for perpetual operation

The amount of energy harvested from the environment can be very different from node to node because of the diversity of harvesters, the locations of the nodes, the deployment policy, the rate of harvesting and so on. In our research we look at solar-based energy harvesting, which has a diurnal characteristic (available readily during the day and no sunlight at night). This actually has an impact on the duty cycle and energy storage dimensioning. In [8] a detailed account of energy harvesting architecture and schemes is presented. Kansal et al. [26] suggested a mathematical condition that would express the conditions under which a sensor node can operate perpetually, through an analysis of the relationship between the harvested energy and the consumed energy as shown in Figure 2.

Suppose $E(t)$ is the power output of any energy source at time $t$, then the energy harvested $E_h$, during any finite time $T$ can be estimated as follows:

$$\rho T - \sigma_1 \leq \int_T E(t)dt \leq \rho T + \sigma_2$$

Therefore, if there is enough historical data, we can make a good estimate of the energy that will be supplied to a node during a given period. Initially when there is no historical data we make use of previous theoretical estimates for the energy harvesting scheme [26].

$E_c(t)$ is the power consumption of a node at a time $t$ and satisfies the following constraint for any value of $T$:

$$\int_T E_c(t)dt \leq \rho' T + \sigma$$

Each node in an energy-harvesting network must have a sensible plan to control the energy that it consumes by means of an energy budget based on adequate historical data. This means that a node must have an energy storage capacity of more than $\sigma + \rho_1 + \rho_2$ and satisfies $\rho' < \rho$ to operate perpetually. Because the energy consumption of each active mode is fixed, the DC of a perpetually operating node completely depends on the calculated value of $\rho'$[9]. The duty cycle is important in low-latency routing for asynchronous WSNs [7]. If a node has a large duty cycle, it can route data more quickly, but it also needs more energy to wake up frequently. Because the quantity of periodically harvested energy is limited, nodes must carefully determine their next hop neighbor and DC.

Figure 2. Energy flow from source to consumer [26].
3.3. Proposed energy consumption model

Dongjin et al. [34] proposed the First Order Radio Model, which has been widely used for measuring energy consumption in wireless communications [29, 35, 36]. In this model, the energy for transmitting one bit of data over distance \( x \) (distance between two nodes) is

\[
E_{tx}(x) = E_{elec} + E_{amp}x^k,
\]

where \( E_{elec} \) is the energy spent by transmitter electronics, \( E_{amp} \) is the transmitting amplifier, and \( k (k \geq 2) \) is the propagation loss exponent. The energy spent by receiver electronics in receiving one bit of data is \( E_{rx} \). The energy consumed to relay one bit data over distance \( x \) is

\[
E_{relay} = E_{elec} + E_{amp}x^k + E_{tx} \equiv E + E_{amp}x^k
\]

(3)

where \( E = E_{elec} + E_{rx} \).

The total threshold energy, \( E_{Th} \) available at each energy harvesting node for transmission at the beginning of each interval is given as:

\[
E_{Th} = E_p + E_h - E_a
\]

where, \( E_p \) is the amount of energy remaining

\( E_h \) is the energy expected from the harvesting device over the next period.

\( E_a \) is an allowance for incorrect assumption.

Some of the threshold energy is used for other energy consuming tasks such as managing the DC, \( E_{dc} \). For the node to be able to send out the information the condition below needs to be satisfied:

\[
E_{relay} + E_{dc} < E_{Th}
\]

3.4. Characteristics of power-adjusted transmission

Stojmenovic and Lin in [36] investigated the characteristics of energy consumption for power adjusted transmission using a generalized form of the First Order Radio Model. Given a source node \( u \) and a destination node \( v \), \( d \) denotes the distance between \( u \) and \( v \), and \( E(d) \) represents the total energy consumed by delivering one bit data from \( u \) to \( v \). The following lemmas hold according to the analysis presented in [17]:

**Lemma 1**

If

\[
d \leq k \sqrt[2-k]{\frac{E}{E_{amp}(1 - 2^{1-k})}}
\]

direct transmission is the most energy-efficient way to deliver packets from \( u \) to \( v \).

**Lemma 2**

If

\[
d > k \sqrt[2-k]{\frac{E}{E_{amp}(1 - 2^{1-k})}}
\]

\( E(d) \) is minimized when all hop distances are equal to \( \frac{d}{d_o} \), and the optimal number of hops is \( \left\lfloor \frac{d}{d_o} \right\rfloor \) or \( \left\lceil \frac{d}{d_o} \right\rceil \), where

\[
d_o = k \sqrt[2-k]{\frac{E}{E_{amp}(k-1)}}
\]
Lemma 3
The total energy consumption for delivering 1 bit data over distance $d$ satisfies $E(d) \geq E \cdot \frac{k}{k-1} \cdot \frac{d}{d_o}$.

It follows from Lemma 2 that $d_o$ is the optimal forwarding distance in terms of minimizing $E(d)$ when $d$ is an integral multiple of $d_o$. Even if $d$ cannot be divided exactly by $d_o$, $d_o$ is a good approximation of the optimal forwarding distance. Moreover, $d_o$ remains a constant for a given sensor device and application environment because $d_o$ only depends on $E_{amp}$, $E$ and $k$. This implies that $d_o$ can be used as an effective metric to enable localized packet forwarding to support energy-efficient routing. On the basis of this observation it follows that the ideal next-hop relay position for any node $u$ (in terms of minimizing the total energy consumption) to deliver a packet from node $u$ to the sink is defined as follows:

Definition 1
Given any node $u$, the ideal position of its next-hop relay, denoted by $f_u$, is defined as the point on the straight line from $u$ to the sink $s$, where $|uf_u| = d_o$.

In our proposed scheme, each node makes fully localized and stateless forwarding decisions based on the location of its ideal next-hop relay position. It locally determines the duty-cycle of each node based on

- Step 1: the estimated energy budget for each period which includes the currently available energy.
- Step 2: the predicted energy consumption (power adjusted transmitter electronics).
- Step 3: the energy expected from the harvesting device. This energy is computed through predictive algorithms using historical data.

4. ENERGY-EFFICIENT BEACONLESS GEOGRAPHIC ROUTING WITH ENERGY SUPPLY

Energy-efficient beaconless geographic routing with energy supply is executed in two modes: beaconless greedy routing and beaconless perimeter recovery mode. EBGRES makes use of three different types of messages, $DATA$, $ACK$, and $SELECT$. Initially the node holding the message/data for forwarding has to determine two relative regions around it because it has information about the destination and its location. These include the positive progress area (PPA) and the negative progress area (NPA). PPA comprises a mode whose position is closest to the destination than the forwarding node (beaconless greedy routing mode). NPA consists of the rest of neighbors of the forwarding nodes that are not providing progress towards the destination (beaconless recovery mode). Additionally, the $DATA$, $ACK$, and $SELECT$ messages include a bit in their header to indicate the routing mode currently being used, that is, Greedy mode ($G$) or Perimeter mode ($P$). The bit is called the routing bit (RB). The RB is set to $G$ mode by default. In beaconless greedy forwarding, only the nodes in the $relay$ $search$ $region$ (Figure 2) of the forwarder are candidates that can forward the packet, and the forwarder chooses the neighbor closest to its optimal relay position as its next-hop relay using the $DATA/ACK$ handshaking mechanism. In this way, each packet is expected to be delivered along the minimum energy route from the source to the sink. If there is no node in the relay search region (i.e., no nodes in the PPA), the forwarder enters the beaconless recovery mode, and the beaconless angular relay is employed to recover from the local minimum.

During the forwarding process we assume that this is true $E_{relay} + E_{dc} < E_{Th}$, to avoid failures. This assumption is based on the fact that the total energy utilized per cycle for relaying data and the energy utilized for duty cycle management must always be less than the threshold energy available or else the node will not be able to forward the data. After sending out the data, the forwarding node is always in receiving mode until it sends out the $SELECT$ message. After that it enters into a sleep or idle mode to save energy while energy is continuously being harnessed from the energy harvesting scheme to maintain the threshold. The same procedure takes place for the receiving node in the relay search region. When a node in the relay search region has been selected the other nodes not participating in the $SELECT$ process will enter into a sleep or idle mode. This duty cycle management is pivotal to facilitate for perpetual operation.
4.1. Relay search region

It has been stated that the best next-hop neighbor for any node \( u \) is the neighbor closest to its ideal relay position \( f_u \), and there is no need for all neighbors of node \( u \) to participate in the contention for acting as the next-hop relay. In EGBRES each node has its relay search region based on the energy available at the power adjusted transmitter, as shown in Figure 3, which is defined as follows:

\[ R_u \]

Given any node \( u \), its next-hop relay search region, denoted by \( R_u \), is defined as the disk centered at \( u \)’s ideal next-hop relay position \( f_u \) with radius \( r_s(u) \) where \( r_s(u) \leq |uf_u| = d_o \).

For any node \( u \), only the neighbors in its relay search region \( R_u \) are candidates that can forward the packets transmitted from node \( u \). The concept of relay search region is introduced to prohibit the unsuitable neighbors from participating in the relay contention procedure.

4.2. Beaconless greedy forwarding algorithm

Given any node \( u \), let \( |us| \) be the distance from node \( u \) to the sink \( s \). Node \( u \) first calculates \( |us| \) because it has the knowledge of its position and the position of the sink. From Lemma 1 it follows that if the sink is in \( u \)’s transmission range and

\[ |us| \leq k \sqrt{\frac{E}{E_{amp}(1 - 2^{1-k})}} \]

node \( u \) transmits its packet directly to the sink because relaying the packets by some intermediate nodes is no more energy-efficient than direct transmission. Otherwise, node \( u \) detects its best next-hop relay node based on the following procedure:

Let \( (x_u, y_u) \) and \( (x_s, y_s) \) be the coordinates of node \( u \) and the sink \( s \), respectively. By Definition 1, the location of \( f_u \), denoted by \( (x_uo, y_uo) \), can be computed as follows:

\[
\begin{align*}
x_{uo} &= x_u - \frac{d_o}{|us|}(x_u - x_s), \\
y_{uo} &= y_u - \frac{d_o}{|us|}(y_u - y_s)
\end{align*}
\]

In the beaconless greedy mode for EGBRES, when the forwarding node \( u \) has a packet to transmit, it sends a broadcast with a \( DATA \) message and waits for responses for a predefined maximum time of \( T_{max} \) seconds.

The \( DATA \) message contains:

- The original message.
- The energy level status.
- The energy harvesting rate.
ENERGY-EFFICIENT BEACONLESS GEOGRAPHIC ROUTING

- The energy consumption rate.
- The duty cycle.
- The position of the forwarding node.
- The location of the ideal next-hop relay position and the radius of its relay search region to detect its best next-hop relay.
- The position of the final destination.

For any neighbor $w$ that receives the DATA message from node $u$, it first determines the relative area in which it is located (PPA or NPA) and then checks if it falls within $R_u$. If $w \neq R_u$, the DATA message is simply discarded. Otherwise, node $w$ generates an acknowledgement (ACK) message, which also contains its own location, identifier, duty cycle, energy status, and sets a proper delay, denoted by $\delta_{w \rightarrow u}$, for broadcasting the ACK message based on a discrete dynamic forward delay function described in the next section. If node $w$ overhears an ACK message that was broadcasted by another candidate before $\delta_{w \rightarrow u}$ expires, node $w$ cancels broadcasting its own ACK message and deletes it (it then enters into sleep or idle mode); otherwise, node $w$ broadcasts its ACK message when $\delta_{w \rightarrow u}$ expires. It is also possible that some neighbors do not receive this ACK message. When node $u$ receives the ACK message (within time, $t \leq T_{\text{max}}$) from its neighbor $w$, the next-hop forwarding node for $u$, denoted by $\text{relay}(u)$, is updated if $\text{relay}(u)$ is null or $|w|_{f_u} < |\text{relay}(u)|_{f_u}$. Finally, node $u$ sends out a SELECT messages to $w$ (its next-hop relay $\text{relay}(u)$). Node $u$ will enter into sleep or idle mode after that confirmation. The header of the messages contains the details of the node selected to be the forwarding node, which in this case is $w$. On receiving the SELECT message the node $w$ sends out the DATA message and the same process is repeated.

4.3. Discrete dynamic forwarding delay

As stated in the previous section all the neighbors wait for a period of time before answering back to the forwarding node. The time to wait is related to their position and this behaviour has two important goals: avoiding collisions and determining the forwarding strategy. Letting all neighbors answer immediately exponentially increases the probability of colliding ACKs because the DATA message arrives almost at the same time as all of them. In EGBRES, the forwarding node selects as nextforwarder the neighbor that is in the relay search region and replies first. Therefore, the forwarding strategy is clearly controlled by timers. By forcing some neighbors to wait more than others we can reduce the number of possible answers and thus, the bandwidth consumption. The goal is to design a function to determine the timeout value in such a way that the most promising neighbors answer in the first place and the other neighbors are suppressed. To determine the forwarding strategy we, as some other protocols in literature, define the progress of a candidate neighbor as a way to measure the goodness of a node as the next forwarder.

For any node $u$, its relay search region $R_u$ is divided into $n$ concentric coronas $S_1, S_2, ..., S_n$ where all coronas have the same area size (Figure 4). $S_i$
Thus, the width of the \( i \)th corona is \((\sqrt{i} - \sqrt{i-1})r_1\), where \( r_1 \) is the radius of \( S_1 \) and \( r_1 = r_s(u)/\sqrt{n} \). If \( v \in S_i \), the distance between node \( v \) and \( f_u \) satisfies that
\[
\frac{\sqrt{i-1} \cdot r_s(u)}{\sqrt{n}} \leq |v| < \frac{\sqrt{i} \cdot r_s(u)}{\sqrt{n}}
\]
Therefore, given any node \( v \in R_u, v \) must locate in \( S_m \)
\[
m = \left[ \left( \frac{\sqrt{m} \cdot |v|}{r_s(u)} \right)^2 \right] + 1
\]
For any node \( v \in R_u, \) instead of broadcasting the \( ACK \) message immediately after receiving the \( DATA \) message from node \( u, \) node \( v \) broadcasts its \( ACK \) with delay \( \delta_{v-u}. \) Let \( \gamma' \) be the delay for transmitting a packet over a unit distance. \( \delta_{v-u} \) is defined as follows:
\[
\delta_{v-u} = \frac{\gamma' \cdot r_s}{\sqrt{n}} \cdot \left( \frac{2^m}{\sqrt{m-1}} + \gamma' \cdot \left( |v| - \sqrt{\frac{m-1}{n}} \cdot r_s(u) \right) \right)
\]
where \( m = \left[ \left( \frac{\sqrt{m} \cdot |v|}{r_s(u)} \right)^2 \right] + 1. \)
The delay computed by (4) guarantees:
1. Nodes in \( S_i \) broadcast their \( ACK \) messages earlier than nodes in \( S_j, \) where \( j > i \)
2. Given node \( v_i \in S_i \) and node \( v_j \in S_j (1 \leq i < j \leq n), \) it must satisfy that \( \delta_{v_j-u} - \delta_{v_i-u} > \gamma' \cdot |v_i| |v_j|, \) which means that \( v_j \) can overhear the \( ACK \) message broadcasted by \( v_i \) and cancels broadcasting its own \( ACK \) message before \( \delta_{v_j-u} \) is due.
3. For all nodes located in the same corona \( S_1, \) the node closest to \( f_u \) broadcasts its \( ACK \) message first because the second term \( \gamma' \cdot \left( |v| - \frac{m-1}{n} r_s(u) \right) \) in (4) guarantees that the node closer to \( f_u \) has a shorter delay.

It follows that for all nodes in the same corona, the node closest to the sink is the best candidate for packet relaying in terms of minimizing source-to-sink energy consumption. In the following, we show that it is not true.

From Figure 4 let \( (0, 0), (25, 0), (21, 2) \) and \( (40, 0) \) be the coordinates for node \( u, a, b, \) and \( s \) respectively. It follows that calculating the distances between the coordinates we obtain
\[
|ub|^2 + |bs|^2 = 21^2 + 2^2 + (40 - 21)^2 + 2^2 = 810
\]
\[
|ua|^2 + |as|^2 = 28^2 + (40 - 25)^2 = 1009
\]
We show that forwarding the packet transmitted by node \( u \) through node \( b \) is more energy-efficient than through node \( a \) although node \( a \) is closer to the sink than node \( b. \) In our scheme, node \( b \) is assigned a shorter delay than node \( a \) and broadcasts the \( ACK \) message first. Thus, node \( b \) becomes the next-hop relay for node \( u \) and received the \( DATA \) message. If there is only one node in the innermost nonempty corona, the above delay function guarantees that the number of \( DATA/ACK \) messages broadcasted for detecting the best relay for \( u \) is minimized (only two messages, one \( DATA \) message broadcasted by \( u \) and one \( ACK \) message broadcasted by the neighbor closest to \( f_u \) ). Even if there are multiple nodes in the innermost nonempty corona, the above delay function can still significantly reduce the number of \( ACK \) messages broadcasted because only the nodes in the most inner nonempty corona have the chance to broadcast \( ACK \) messages. This leads to a reduction in the number of collisions, wastage of energy and network bandwidth. On the basis of this delay function, each node can simply use the largest relay search region (i.e., \( r_s(u) = d_o \)) to cover more candidates because the delay function can effectively suppress unsuitable neighbors in the relay search region to broadcast \( ACK \) messages. When the \( ACK \) is received the forwarding node \( u \) then stops the timer and sends a \( SELECT \) message to the new forwarding node.
4.4. Beaconless recovery (perimeter routing)

When node \( u \) broadcasts a DATA message to forward a message to its best next-hop relay, it sets its timer to \( T_{\text{max}} \) and starts the timer. \( T_{\text{max}} \) is large enough to guarantee that node \( u \) can receive the ACK message from the farthest neighbor in \( R_u \) before the timer expires as explained above when calculating the DDFD. If node \( u \) receives no ACK message till the timer expires, it assumes that there is no neighbor in its relay search region or the packet has been lost or the data has been sent to the NPA region. In this case, a so-called void area is found and the routing process cannot continue in the greedy mode. In geographic routing protocols we have different strategies to enclose these void areas, but it is necessary to know the position of the neighbors to locally build a planar graph to determine the next perimeter forwarder. Common algorithms to do that include the RNG and GG [23]. Face routing requires the network topology to be a planar graph (i.e., no edges intersect each other). To planarize a graph, several algorithms can be used [23]. GG, for instance, contains edges between nodes \( u \) and \( v \) if and only if no other nodes are located inside the circle centered in the middle of edge \((u, v)\) and with diameter \(|uv|\). GG has some desirable properties when used for routing in wireless networks such as localized message, free computation, planarity, and preserving connectivity [1]. GG divides the network into faces. In RNG a node \( u \) preserved all outgoing edges \( uv \) which satisfies that the intersection of the circles with center, \( u \), center \( v \), and radii \( uv \) contains no other node than \( v \).

As we have mentioned previously, in EGBRES, the ACK messages from neighbors in the NPA do not cancel any timer. Thus, when the forwarding node does not have any neighbor that can be used as a relay node towards the destination, all the ACK messages from the other neighbors (those in the NPA area) are received and stored. When the timer expires, the forwarding node can thus build the planar graph using all the information gathered and then selects the next forwarding node using the desired recovery mechanism. In our case we use the GFG algorithm where greedy routing is the same as in [5] while Face routing is similar to the one in [23]. One of the drawbacks of Face routing is that it is likely to follow short edges of GG that may be power inefficient.

In the case of perimeter routing the SELECT message must include some extra information which includes the identifier of the forwarding node, its position, the identifier of the next hop selected and the current perimeter information defined by GFG which consists of:

- The position of the node where perimeter routing started \( L_p \),
- The first edge (\( E_0 \)) traversed on current face, and
- The \( (L_p') \) point, that is the cross point between the \( L_p s' \) line and the current face, being \( s' \) the position of the destination node.

Additionally, the RB bit must be set to \( P \) to indicate perimeter routing. When messages are being routed in perimeter routing, the behaviour of neighbors is slightly different. To begin with, the forwarding node must include in the DATA message the \( L_p \) point. A neighbor receiving the DATA message must check if it is placed closer to the destination than the \( L_p \) point. If that is the case, it switches from perimeter routing to greedy mode. Thus, the RB bit of the ACK message is set to \( G \). Only in those cases, the ACK message will be sent but only after the timer expires (as in greedy routing). The forwarding node may stop its timer if it receives an ACK message with \( RB \equiv G \). In that case, it switches to greedy mode and continues in that mode by sending the appropriate SELECT message. If there is no neighbor closer to the destination than \( L_p \) then the forwarding node will wait up to \( T_{\text{max}} \) seconds. Then, it selects the next forwarder to continue in perimeter mode and selects it by broadcasting the corresponding SELECT message with \( RB \equiv P \).

5. THEORETICAL ANALYSIS OF EGBRES

In this section we present a theoretical analysis for EGBRES. This analysis is initially based on a simplified Media Access Control (MAC) model without packet loss, the unit disk graph model without failures in greedy forwarding and uniform node deployment. We will analyze the algorithm with packet loss and failure in greedy forwarding. We also look at the selection of the relay node in the relay search region, which depends on the energy available at any time, the status
of the node (assumed to be in receiving mode), and the fact that we use a power adjusted transmitter where energy is always nearly at its maximum. We investigate the effect of increased and consistent energy supply across all nodes of the wireless sensor network. We will present design considerations in a realistic scenario for the proposed EBGRES algorithm.

5.1. Definitions and notations used in theoretical analysis

Melodia et al. [37] defined two terms, progress and advance, which were introduced to distinguish different forwarding rules in geographic routing. Suppose that node $u$ forwards its packets to its neighbor $v$ for relay to the sink $s$. The progress, denoted by $P(u,v)$, is defined as the projected distance of $|uv|$ on the straight line passing through $u$ and $s$, and the advance, denoted by $A(u,v)$, is defined as the difference between $|us|$ and $|vs|$. Therefore,

$$P(u,v) = |uv| \cos \angle vus$$

$$A(u,v) = |us| - |vs|$$

We define two metrics called energy over progress ratio and energy over advance ratio to analyze the characteristics of energy consumption in EBGRES. Let $\gamma_P(u,v)$ and $\gamma_A(u,v)$ be the energy over
progress ratio and the energy over advance ratio for relaying 1 bit data from \( u \) to, \( v \) respectively. \( \gamma_P(u, v) \) and \( \gamma_A(u, v) \) are defined as follows:

\[
\gamma_P(u, v) = \frac{E_{\text{relay}}}{P(u, v)} = \frac{E + E_{\text{amp}}|uv|^k}{|uv|\cos \angle vu \nu s} \tag{7}
\]

\[
\gamma_A(u, v) = \frac{E_{\text{relay}}}{A(u, v)} = \frac{E + E_{\text{amp}}|uv|^k}{|us| - |vs|} \tag{8}
\]

Without loss of generality, we assume that the maximum transmission range \( R \) is no less than \( 2d_o \) because the analysis approach for \( R < 2d_o \) is the same.

5.2. Guaranteed delivery

**Theorem 1**

EBGRES is loop-free in greedy forwarding mode

**Proof**

Let \( u \) be a source node and \( s \) be the sink. If \( s \) locates in the transmission range of node \( u \) and

\[
|us| \leq \sqrt[k]{\frac{E}{E_{\text{amp}}(1 - 2^{1-k})}},
\]

node \( u \) sends its packets directly to the sink without any relay. The theorem holds in this case.

If

\[
|us| > \sqrt[k]{\frac{E}{E_{\text{amp}}(1 - 2^{1-k})}}
\]

the packets generated by node \( u \) may be relayed by some intermediate nodes before arriving at the sink. As shown in Figure 5, the maximum distance from \( s \) to any point in \( R_u \) is \( |sb| \).

\( \forall v_1 \in R_u, |v_1s| \leq |bs| < |us| \) because \( r_s(u) \leq d_o \) and no two nodes are located at the same position. Therefore, \( A(u, v_1) = |us| - |v_1s| > 0 \), which means that each forwarding must obtain a positive advance.

Let \( v_n \) be a node that relays the packets generated by node \( u \), and \( v_0v_1 \ldots v_m \ldots v_{n-1}v_n \) represent the routing path from \( u \) to \( v_n \) in EBGRES where \( v_0 = u \). For any relay node \( v_m \) prior to \( v_n \) in this routing path, \( A(v_n, v_m) = |v_ns| - |v_ms| < 0 \), which means that \( v_n \) cannot forward its packets to \( v_n \). Hence, the theorem holds.

By Theorem 1, EBGRES is loop-free in beaconless greedy forwarding mode. In beaconless recovery mode (Perimeter routing), the beaconless angular relaying uses the Select and Protest methods to
avoid crossing edges which might cause a routing loop. In [24], it is proven that the angular relaying algorithm can always select the first edge of the Gabriel subgraph in counterclockwise order. Thus, there are no routing loops in angular relaying. Therefore, EBGRES can provide guaranteed delivery as long as the network is connected.

5.3. Bounds of hop count

For any node \( u \), let \( C(u) \) be the minimum relay search region that covers only one neighbor, and let \( r \) be the radius of \( C(u) \). Because nodes are uniformly deployed, the minimum relay search regions for all nodes in the network have the same size. Then, we have the following theorem:

**Theorem 2**

If there are no failures in greedy forwarding, the number of hops, denoted by \( N \), for delivering a packet from \( u \) to the sink \( s \) under EBGRES where \( |us| = d \) satisfies

\[
\frac{d}{d_o + r} - 1 < N < \frac{d}{d_o - r} + 1
\]

**Proof**

Let \( v \) be the node in \( C(u) \). As shown in Figure 6, \( |vs| \) is maximized when \( v \) is located at point \( a \). By (6), \( A(u, v) = |us| - |vs| \). Therefore, \( d_o - r \leq A(u, v) \leq d_o + r \)

Let \( u_1, u_2, ..., u_{N-1}s \) denote the routing path from \( u \) to the sink \( s \). On the basis of (6),

\[
d = |us| = A(u, v_1) + |v_1s|
\]

\[
= A(u, v_1) + A(v_1, v_2) + |v_2s|
\]

\[
= A(u, v_1) + A(v_1, v_2) + ... + A(v_{N-2}, v_{N-1}) + |v_{N-1}s|
\]

because \( |v_{N-1}s| = A(v_{N-1}, s) \),

\[
d = A(u, v_1) + \sum_{i=1}^{N-2} A(v_i, v_{i+1}) + A(v_{N-1}, s)
\]

For the prior \( N - 1 \) hops, the best relay is chosen based on the same metric (i.e., the neighbor closest to the ideal next-hop relay position). Therefore,

\[
(N - 1)(d_o - r) \leq d - A(v_{N-1}, s) \leq (N - 1)(d_o + r)
\]

That is,

\[
\frac{d - A(v_{N-1}, s) + (d_o + r)}{d_o + r} \leq N \leq \frac{d - A(v_{N-1}, s) + (d_o - r)}{d_o - r}
\]

![Figure 6. C(u) is the minimum relay search region that covers only one node and r is the radius of C(u).](image)
For the last hop, the packet is directly transmitted to the sink. On the basis of Lemma 1,

\[ 0 < A(v_{N-1}, s) \leq k \sqrt{\frac{E}{E_{\text{amp}}(1 - 2^{1-k})}} \]

**Lemma 8**

\[ d_o < k \sqrt{\frac{E}{E_{\text{amp}}(1 - 2^{1-k})}} < 2d_o, \text{ when } k \geq 2 \]

**Proof**

Let

\[ f(k) = k \sqrt{\frac{E}{E_{\text{amp}}(1 - 2^{1-k})}} = k \sqrt{\frac{k-1}{(1 - 2^{1-k})}} = k \sqrt{\frac{k-1}{2^k - 2}} \]

When \( k \geq 2 \), \( 1 - 2^{1-k} < k-1 < 2^k - 2 \). Thus \( 1 < f(k) < k\sqrt{2^k} = 2 \). By Lemma 1,

\[ d_o = k \sqrt{\frac{E}{E_{\text{amp}}(k-1)}}. \]

Therefore

\[ d_o < k \sqrt{\frac{E}{E_{\text{amp}}(1 - 2^{1-k})}} < 2d_o \]

By Lemma 8 it follows that:

\[ 0 < A(v_{N-1}, s) < 2d_o < 2(d_o + r) \]

Thus,

\[ N \geq \frac{d - A(v_{N-1}, s) + (d_o + r)}{d_o + r} > \frac{d - 2(d_o + r) + (d_o + r)}{d_o + r} = \frac{d}{d_o + r} - 1 \]

and

\[ N \leq \frac{d - A(v_{N-1}, s) + (d_o - r)}{d_o - r} < \frac{d + (d_o - r)}{d_o - r} = \frac{d}{d_o - r} + 1 \]

Therefore,

\[ \frac{d}{d_o + r} - 1 < N < \frac{d}{d_o - r} + 1 \]

**5.4. Upper bound on energy consumption**

In EBGRES, the best next-hop relay for each node is detected through \( \text{DATA}/\text{ACK} \) handshaking. Because the discrete delay function can effectively suppress unsuitable candidates from broadcasting \( \text{ACK} \) messages, the number of \( \text{DATA}/\text{ACK} \) messages broadcasted is proportional to the number of hops for delivering the data to the sink. Therefore, the energy consumed by broadcasting and receiving \( \text{DATA}/\text{ACK} \) can be viewed as a part of energy spent by data delivery. In this analysis, the energy consumption for source-to-sink data delivery is referred to as the sum of energy consumed by the nodes in the routing path for delivering 1 bit data from the source to the sink. We first prove that the position that maximizes the energy over advance ratio \( \gamma_A(u, v) \) must be located on the border of the minimum relay search region in Lemma 4. Then, in Lemma 6, we demonstrate that the energy over advance ratio at each hop has an upper bound. Finally, the upper bound on energy...
consumption for source-to-sink data delivery is given in Theorem 3. The availability of perpetual energy across all nodes has an impact of improving packet delivery, prolonging the network lifetime and workload. This will not affect the upper bound of energy consumption but offers more energy which can be utilized by the nodes for transmission and sensing of the data.

**Lemma 4**
Given $v \in C(u)$, $v$ must be located on the border of $C(u)$ when $\gamma_A(u, v)$ is maximized.

**Proof**
Suppose $v$ is located inside $C(u)$. Let $v'$ be a point on the border of $C(u)$ and $v'$ has the same $x$-coordinates with $v$. By Equation (8),

$$\gamma_A(u, v') - \gamma_A(u, v) = \frac{E_{\text{relay}}(|uv'|)}{\text{A}(u, v')} - \frac{E_{\text{relay}}(|uv|)}{\text{A}(u, v)}$$

$$= \frac{E + E_{\text{amp}}(|uv'|^k)}{|us| - |v's|} - \frac{E + E_{\text{amp}}(|uv|^k)}{|us| - |vs|}$$

Because $|v's| > |vs|$ and $|uv'| > |uv|$, $\gamma_A(u, v') - \gamma_A(u, v) > 0$. Hence, this Lemma holds. □

**Lemma 5**
If $v$ is located on the border of $C(u)$ and $r > 0$,

$$\frac{A(u, v)}{P(u, v)} \geq \frac{r}{d_o + r - \sqrt{d_o^2 - r^2}}$$

**Proof**
As shown in Figure 7, $|vs| = |ns|$ and $|vd| = |md|$, Clearly, $|un| \geq |um|$ because $\angle vns > \angle vms$. By (7) and (8),

$$\frac{A(u, v)}{P(u, v)} = \frac{|us| - |v|}{ue} = \frac{|us| - |ns|}{ue} \geq \frac{|ud| - |md|}{ue} = \frac{|ud| - |vd|}{ue}$$

Assume that $u$ is the origin and the straight line from $u$ to $s$ is the axis. $(0, 0)$ and $(x, y)$ represent the coordinates of $u$ and $v$, respectively. Because $v$ is located on the border of $C(u)$, $(x-d_o)^2 + y^2 = r^2$. Thus,

![Figure 7](image_url)

Figure 7. $C(u)$ is the minimum relay search region and $v$ is a node located on the border of $C(u)$, where $|vs| = |ns|$, $|vd| = |md|$, and $|ud| = |ub|$. 

\[
\frac{A(u, v)}{P(u, v)} \geq \frac{d_o + r - \sqrt{2r(r + d_o - x)}}{x} \\
\geq \frac{r}{d_o + r - \sqrt{d_o^2 - r^2}}
\]

\[\square\]

Lemma 6

\( v \in C(u), \gamma_A(u, v) < \max \left\{ \frac{2E_{\text{amp}}d_o^k}{(d_o - r)(1 + \sqrt{d_o^2 - r^2})}, \frac{2E_{\text{amp}}(k - 1)d_o^k + (d_o + r)^k}{(d_o + r)(\sqrt{d_o^2 - r^2} + d_o - r)} \right\} \]

Proof

From Lemma 4, \( v \) must be located on the border of \( C(u) \) when \( \gamma_A(u, v) \) is maximized. By (8) and Lemma 5,

\[\gamma_A(u, v) = \frac{E + E_{\text{amp}}(|uv|^k)}{A(u, v)} = \frac{E + E_{\text{amp}}(|uv|^k)}{P(u, v) \cdot \frac{A(u, v)}{P(u, v)}} = \gamma_A(u, v) \cdot \frac{P(u, v)}{A(u, v)}\]

Thus,

\[\gamma_A(u, v) \leq \gamma_P(u, v) \frac{d_o + r - \sqrt{d_o^2 - r^2}}{r}\]

Let \( \overline{ab} \) be the tangent of the circle centered at \( f_u \) with radius \( r \) (see Figure 6). \( a \) and \( b \) are two points on the line segment from \( a \) to \( b \) where \(|ua| = (d_o - r) \sec \theta \) and \(|ub| = d_o + r \). Let \( \psi \) denote the set of points located on the line segment between \( a \) and \( b \). \( \forall w \in \psi \),

\[\gamma_A(u, w) = \frac{E + E_{\text{amp}}(|uw|^k)}{P(u, v)} = \frac{1}{\cos \theta} \cdot \frac{E + E_{\text{amp}}(|uw|^k)}{|uw|}\]

where \((d_o - r) \sec \theta \leq |uw| \leq d_o + r\). Clearly, \( \gamma_P(u, w) \) given in (10) is strictly concave with respect to \(|uw|\). Therefore,

\[\max_{w \in \psi} \gamma_P(u, w) = \max\{\gamma_P(u, a), \gamma_P(u, b)\}\]

Similar to the proof of Lemma 4, we can prove that \( \gamma_P(u, w) \leq \max_{w \in \psi} \gamma_P(u, w) \). By Lemma 5 and (10),

\[\gamma_A(u, w) < \max\{\gamma_P(u, a), \gamma_P(u, b)\} \cdot \frac{d_o + r - \sqrt{d_o^2 - r^2}}{r}\]

where

\[\gamma_A(u, a) = \frac{E + E_{\text{amp}}d_o^k \left( \frac{d_o - r}{d_o + r} \right)^\frac{k}{2}}{d_o - r}\]

and

\[\gamma_A(u, b) = \frac{d_o(E + E_{\text{amp}}(d_o + r)^k)}{(d_o + r)\sqrt{d_o^2 - r^2}}\]

By Lemma 2, \( d_o = \sqrt[2k]{\frac{E_{\text{amp}}}{E_{\text{amp}}(k - 1)}} \). Replacing \( \gamma_A(u, a) \) and \( \gamma_A(u, b) \) in (12),
\[
\gamma_A(u, v) = \frac{2E_{\text{amp}}d_o^k \left[ k - 1 + \left( \frac{d_o - r}{d_o + r} \right) \right]}{(d_o - r) \left( 1 + \sqrt{\frac{d_o - r}{d_o + r}} \right)} \cdot \frac{2E_{\text{amp}}d_o \left[ (k - 1)d_o^k + (d_o + r)^k \right]}{(d_o + r) \left( \sqrt{d_o^2 - r^2} + d_o - r \right)}
\]

**Theorem 3**
If there are no failures in greedy forwarding and no packet loss in EBGRES, the total energy consumption, denoted by \( E(d) \), for delivering 1 bit data from source \( u \) to the sink \( s \) under EBGRES satisfies

\[
E(d) < \max \left\{ \frac{2E_{\text{amp}}d_o^k \left[ k - 1 + \left( \frac{d_o - r}{d_o + r} \right) \right]}{(d_o - r) \left( 1 + \sqrt{\frac{d_o - r}{d_o + r}} \right)} \cdot \frac{2E_{\text{amp}}d_o \left[ (k - 1)d_o^k + (d_o + r)^k \right]}{(d_o + r) \left( \sqrt{d_o^2 - r^2} + d_o - r \right)} \right\} \cdot d
\]

**Proof**
Let \( E'(d) \) and \( N' \) be the total energy consumption and the number of hops, respectively, for delivering 1 bit data for \( u \) to the sink \( s \) when all hops including the last one use the same metric (i.e., the neighbor closest to the ideal next-hop relay position) to choose a relay node. Let \( d_i \) denote the distance for the \( i \)th hop. Let \( A_i \) and \( \gamma(i) \) denote the advance and the energy over advance ratio for the \( i \)th hop, respectively:

\[
E'(d) = \sum_{i=1}^{N'} E_{\text{elec}} + E_{\text{amp}}d_1^k + \sum_{i=2}^{N'} E_{\text{relay}}(d_i) + \frac{E_{\text{rx}}}{s} = \sum_{i=1}^{N'} E_{\text{relay}}(d_i) = \sum_{i=1}^{N'} \gamma(i) A_i
\]

Because all hops are independent and use the same routing metric, \( E'(d) \) is maximized when the packet is forwarded to the neighbor with the maximum energy over advance ratio at each forwarding node. By Lemma 1, \( E(d) < E'(d) \) because the metric used for the last forwarding in EBGRES saves energy. Therefore,

\[
E(d) < E'(d) < \left( \max_{i=1}^{N'} \gamma(i) \right) \cdot \sum_{i=1}^{N'} A_i
\]

On the basis of the proof of Theorem 2, \( \sum_{i=1}^{N'} A_i = d \). Thus

\[
E(d) < d \cdot \max_{i=1}^{N'} \gamma(i)
\]

By Lemma 6, the upper bound on energy over advance ratio only depends on \( E_{\text{amp}}, d_o, k \) and \( r \). Therefore, all hops have the same upper bound on energy over advance ratio. Therefore,

\[
E(d) < \max \left\{ \frac{2E_{\text{amp}}d_o^k \left[ k - 1 + \left( \frac{d_o - r}{d_o + r} \right) \right]}{(d_o - r) \left( 1 + \sqrt{\frac{d_o - r}{d_o + r}} \right)} \cdot \frac{2E_{\text{amp}}d_o \left[ (k - 1)d_o^k + (d_o + r)^k \right]}{(d_o + r) \left( \sqrt{d_o^2 - r^2} + d_o - r \right)} \right\} \cdot d
\]

Let \( r_{ul} \) denote the ratio of the upper bound to the lower bound on energy consumption for delivering 1 bit data over distance \( d \) in EBGRES. We have the following corollary:
Corollary 1

\[
R_{ul} < \min \left\{ \frac{k - 1 + \left( \frac{d_x - r}{d_y + r} \right)^{\frac{k}{2}}}{k \left( 1 - \frac{r}{d_y} \right) \sqrt{\frac{d_x - r}{d_y + r}}} \right\} \cdot \frac{1}{E_{ul}} \cdot \frac{d}{d_y} \cdot (d_x - r) \left( 1 + \frac{d_x - r}{d_y + r} \right) \cdot d
\]

when \( E_{rx} = E_{elec} = 50 \text{ nJ/bit}, E_{amp} = 100 \text{ nJ/bit/m}^2 \) and \( k = 2, R_{ul} < \frac{1000}{\sqrt{1000 - r}(\sqrt{1000} - r^2)} \).

Proof

By Lemma 3 and Theorem 3, we have

\[
R_{ul} = \frac{E_{upperbound}(d)}{E_{lowerbound}(d)}
\]

\[
< \min \left\{ \frac{2E_{amp}d_x^k [k - 1 + \left( \frac{d_x - r}{d_y + r} \right)^{\frac{k}{2}}]}{(d_x - r)(1 + \frac{d_x - r}{d_y + r})} \cdot \frac{d}{d_y} \cdot (d_x - r) \left( 1 + \sqrt{\frac{d_x - r}{d_y + r}} \right) \cdot \frac{d}{d_y} \cdot (d_x + r) \left( \sqrt{\frac{d_x - r}{d_y + r}} + d_x - r \right) \right\}
\]

By Lemma 2, \( d_o = \sqrt[4]{\frac{E}{E_{amp}(k-1)}} \). Replacing \( d_o \) with \( \sqrt[4]{\frac{E}{E_{amp}(k-1)}} \)

\[
R_{ul} < \min \left\{ \frac{k - 1 + \left( \frac{d_x - r}{d_y + r} \right)^{\frac{k}{2}}}{\frac{1}{2} \sqrt{\frac{d_x - r}{d_y + r}} \left( 1 - \frac{r}{d_y} + \sqrt{1 - \frac{r^2}{d_y^2}} \right)} \right\} \cdot \frac{1}{2} \left( 1 + \frac{r}{d_y} \right) \left( 1 - \frac{r}{d_y} + \sqrt{1 - \frac{r^2}{d_y^2}} \right)
\]

Because \( 1 - \frac{r}{d_y} < \sqrt{1 - \frac{r^2}{d_y^2}} \)

\[
R_{ul} < \min \left\{ \frac{k - 1 + \left( \frac{d_x - r}{d_y + r} \right)^{\frac{k}{2}}}{k \left( 1 - \frac{r}{d_y} \right) \sqrt{\frac{d_x - r}{d_y + r}}} \right\} \cdot \frac{1}{2} \left( 1 + \frac{r}{d_y} \right) \left( 1 - \frac{r}{d_y} + \sqrt{1 - \frac{r^2}{d_y^2}} \right)
\]

Similar to [34, 38], the system parameters are set as follows: \( E_{rx} = E_{elec} = 50 \text{ nJ/bit}, E_{amp} = 100 \text{ nJ/bit/m}^2 \) and \( k = 2 \), by Lemma 2, \( d_o = \sqrt[4]{\frac{E}{E_{amp}}} = \sqrt[4]{10,000} \). Thus

\[
R_{ul} < \frac{1000}{\sqrt{1000 - r}(\sqrt{1000} - r^2)}.
\]

From Corollary 1, it is worth noting that \( R_{ul} \rightarrow 1 \) when \( r \rightarrow 0 \). However, \( R_{ul} \) becomes infinite when approaching \( d_o \). This phenomenon can be explained as follows: when \( r \) approaches to \( d_o \), as can be seen in Figure 5, the advance obtained by forwarding the data to the next-hop relay is very small if the next-hop relay is located at the position closest to node \( u \). Thus, \( y_A(v, u) \) must be very large because the energy spent by the electronic circuit (i.e., \( E_{elec} \) and \( E_{rx} \)) becomes the dominant part in the total energy consumption. The worst case is when the next-hop relay for each forwarding node is located at the position that maximizes the energy over advance ratio. Let \( Prob_o(d) \) denote the
probability that the worst case happens for delivering packets from \(u\) to the sink \(s\) where \(|us| = d\).

The following theorem shows that \(\text{Prob}_{\omega}(d)\) approaches to 0 when \(r\) approaches to \(d_o\):

**Theorem 4**

\(\text{Prob}_{\omega}(d)\) monotonically decreases with the increase of \(d\) and \(r\). \(\text{Prob}_{\omega}(d) \rightarrow 0\) when \(r_{ul} \rightarrow d_o\).

**Proof**

Let \(N\) be the number of hops for delivering a packet from \(u\) to \(s\) where \(|us| = d\). Let \(p(i)\) denote the probability that the packet is forwarded in the way that the energy over advance ratio at hop \(i\) is maximized. For the prior \(N - 1\) hops, the forwarding at each hop is the same and independent of the number of nodes. Thus, \(p(i) = p(j) = p(1 \leq i, j \leq N - 1)\). We have

\[
\text{Prob}_{\omega}(d) = p(N) \cdot \prod_{i=1}^{N-1} p(i) < p^{N-1}
\]

Because nodes are deployed with uniform distribution within this analysis, \(pa(1/\pi r^2)\). For the hop count \(N\), it increases when \(d\) increases. Therefore, \(\text{Prob}_{\omega}(d)\) monotonically decreases when \(d\) and \(r\) increase. From (5), the location that maximizes \(\gamma_A(u, v)\) approaches to \(u\) when \(r \rightarrow d_o\) because the advance obtained with each forwarding step is very small and \(E\) becomes the dominant part in energy consumption. When \(r \rightarrow d_o, N \rightarrow \infty\). Therefore, \(\text{Prob}_{\omega}(d) \rightarrow 0\) when \(r \rightarrow d_o\).

**5.5. Expected energy consumption**

Let \(E[\gamma_A(u, v)]\) denote the expected energy over advance ratio for one hop forwarding in EBGRES. We have the following lemma:

**Lemma 7**

\[
E[\gamma_A(u, v)] \approx \rho \int \int_{C(u)} \frac{E + E_{\text{amp}}(x^2 + y^2)}{x} \, dx \, dy
\]  

where \(\rho\) is the node density.

**Proof**

Let \(v\) be the node in \(C(u)\). Because sensor nodes are uniformly distributed with density \(\rho\), by (8),

\[
E[\gamma_A(u, v)] = \int \int_{C(u)} \frac{E_{\text{relay}}(|uv|)}{A(u, v)} \, dx \, dy
= \rho \int \int_{C(u)} \frac{E + E_{\text{amp}} |uv|^k}{|us| - |vx|} \, dx \, dy
\]

![Figure 8. Approximation of advance where \(v\) is the only node in \(C(u)\) and \(|vx| = |bs|\).](image)
When nodes are densely deployed, $C(u)$ is small. As shown in Figure 8, the advance obtained by forwarding the packet to $v$ is close to the progress, that is, $|us| - |vs| \approx |ua|$. Thus,

$$E[\gamma_A(u, v)] \approx \rho \iint_{C(u)} \frac{E + E_{\text{amp}}|uv|^k}{|ua|} \, dx \, dy$$

$$= \rho \iint_{C(u)} \frac{E + E_{\text{amp}}(x^2 + y^2)^{k/2}}{x} \, dx \, dy \quad (14)$$

**Theorem 5**

If there are no failures in greedy forwarding and no packet loss, the expected energy consumption, denoted by $E[E(d)]$ for delivering 1 bit data from source $u$ to the sink $s$ where $|us| = d$ satisfies

$$E[E(d)] \approx \rho d \iint_{C(u)} \frac{E + E_{\text{amp}}(x^2 + y^2)^{k}}{x} \, dx \, dy$$

**Proof**

Let $N$ be the number of hops to deliver one packet from $u$ to $s$. By Lemma 7, the approximation of $E[\gamma_A(u, v)]$ only depends on $\rho$. Therefore, the prior $N - 1$ hops have the same approximated energy over advance ratio, denoted by $E[\gamma|]$, because of the same forwarding procedure. Let $E[\gamma(N)]$ be the energy over advance ratio for the last hop. As shown in Figure 8,

$$E[\gamma(N)] \approx \rho \iint_{C(u)} \frac{E + E_{\text{amp}}|sv|^k}{|sv|} \, dx \, dy$$

When nodes are densely deployed, $C(u)$ is small and $E[\gamma(N)] \approx E[\gamma]$. Furthermore, the effect of the last hop on the total energy consumption is small for large $d$. Therefore,

$$E[E(d)] = E \left[ \sum_{i=1}^{N} \gamma(i) d_i \right] \approx E[\gamma] d$$

$$\approx \rho d \iint_{C(u)} \frac{E + E_{\text{amp}}(x^2 + y^2)^{k}}{x} \, dx \, dy$$

Let $r_{el}$ be the ratio of the approximated expected energy consumption to the lower bound on energy consumption for delivering 1 bit data from a node to the sink. We have the following corollary:

**Corollary 2**

$$r_{el} = \frac{1}{2} + \frac{12d_o^2 + \pi r^2}{24\sqrt{\pi r}d_o} \ln \frac{2d_o + \sqrt{\pi r}}{2d_o - \sqrt{\pi r}} \text{ when } k = 2$$

and $r_{el} < 1.5$ when $E_{\text{elec}} = E_{rx} = 50 \text{ nJ/bit}$ and $E_{\text{amp}} = 100 \text{ pJ/bit/m}^2$.

**Proof**

As shown in Figure 7, a square centered at $f_u$ with side $l$ is used to approximate $C(u)$ where $l^2 = \pi r^2$. Because $C(u)$ is the minimum relay search region, $\rho \pi r^2 = 1$, that is $\rho = \frac{1}{\pi r^2}$. When $k = 2$,

$$E[E(d)] \approx \frac{d}{\pi r^2} \int_{d_o-1}^{d_o+1} \int_{-l}^{l} \frac{E + E_{\text{amp}}(x^2 + y^2)^{k/2}}{x} \, dx \, dy$$

$$\approx \frac{E_{\text{amp}}d(12d_o^2 + \pi r)}{12\sqrt{\pi r}} \ln \frac{2d_o + \sqrt{\pi r}}{2d_o - \sqrt{\pi r}} + E_{\text{amp}} d d_o \quad (15)$$
By (15) and lemma 3,
\[
    r_{el} = \frac{1}{2} + \frac{12d_o^2 + \pi r^2}{24\sqrt{\pi}rd_o} \ln \frac{2d_o + \sqrt{\pi r}}{2d_o - \sqrt{\pi r}}
\]
where \( E_{elec} = E_{rx} = 50 \text{ nJ/bit}, E_{amp} = 100 \text{ pJ/bit/m}^2 \) and \( k = 2 \), by Lemma 2, \( d_o = \sqrt{\frac{E}{E_{amp}}} = \sqrt{1000} \). When \( r \to 0 \), \( r_{el} \to 1 \). When \( r = \sqrt{1000} \), \( r_{el} < 1.5 \). □

5.6. Summary of analysis on energy consumption

To demonstrate the energy efficiency of EBGRES, we present the comparison between \( r_{ul} \) and \( r_{el} \). Similar to [39, 40], the system parameters are set as follows: \( k = 2 \), \( E_{elec} = E_{rx} = 50 \text{ nJ/bit} \), and \( E_{amp} = 100 \text{ pJ/bit/m}^2 \). Under this setting, the optimal hop distance is \( \sqrt{1000} \) according to Lemma 2. The maximum transmission range \( R \) for all nodes is set to 80 m which is larger than \( 2d_o \). Figure 9 plots the ratio \( r_{ul} \) and \( r_{el} \) under different sizes of the relay search region. \( r_{el} \) and \( r_{ul} \) are the ratio of the approximated expected energy consumption to the lower and upper bound, respectively, on energy consumption for delivering 1 bit data from a node to the sink. We observe that \( r_{ul} \) is close to 1 when \( r \) is small because a small \( r \) means that the best relay for each node keeps very close to its ideal next-hop relay position. As \( r \) increases, \( r_{ul} \) first increases slightly and reaches around 2 when \( r = 15 \text{ m} \). After that, \( r_{ul} \) increases quickly with the increase of \( r \) and approaches to infinity when \( r \) comes near to \( \sqrt{1000} \text{ m} \). However, \( r_{el} \) increases slowly with the increase of \( r \). Even if \( r = \sqrt{1000} \), \( r_{el} \) is less than 1.5. This is because the probability that packets are delivered along the worst path decreases as \( r \) increases.

5.7. Routing metric for unreliable communication links

The protocol design and theoretical analysis discussed in the previous sections are based on the unit disk graph model in which transmissions between any two nodes within the communication range are assumed to be reliable. In this section, we extend EBGRES to lossy sensor networks to provide energy efficient source-to-sink routing in the presence of unreliable communication links.

To capture the characteristics of data loss, packet reception rate (PRR) is used to measure the quality of communication links. Let \( PRR(u, v) \) be the packet reception rate for link \((u, v)\). \( PRR(u, v) \) is defined as the ratio of the number of successful transmissions from \( u \) to \( v \) to the total number of transmissions from \( u \) to \( v \). Thus, the expected number of transmissions that guarantees one successful transmission from \( u \) to \( v \) is \( \frac{1}{PRR(u, v)} \).

![Figure 9. Comparison of \( r_{ul} \) and \( r_{el} \) with the variation of the radius of the minimum relay search region.](image-url)
Packets may be lost because of many reasons such as data corruption, collision, or the attenuation of signal strength. In the case where a packet is lost before reaching the receiver, nearly the same amount of energy is dissipated by listening [31]. Therefore, the total expected energy consumption for successfully relaying 1 bit data from \( u \) to \( v \) denoted by \( E[E(u, v)] \), can be approximately modeled as

\[
E[E(u, v)] \approx \frac{E_{\text{relay}}(|uv|)}{PRR(u, v)}
\]

By (8), the expected energy over advance ratio, denoted by \( E[E(u, v)] \), for successfully relaying 1 bit data from \( u \) to \( v \), satisfies

\[
E[E(u, v)] \approx \frac{E_{\text{relay}}(|uv|)}{PRR(u, v)A(u, v)} \approx \frac{1}{PRR(u, v)} \cdot \frac{E_{\text{relay}}(|uv|)}{A(u, v)}
\]

(16)

Note that the second part \((E_{\text{relay}}(|uv|))/(A(u, v))\) in (16) is the energy over advance ratio when transmission from \( u \) to \( v \) is reliable. Motivated by this observation, we propose a new metric for providing energy-efficient routing in lossy sensor networks. Instead of choosing the neighbor closest to the ideal next-hop relay position among all candidates in the relay search region, node \( u \) chooses the neighbor that minimizes \( |vf_u|/PRR(u, v) \) as its next-hop relay. We refer to the extended version of EBGRES using this routing metric as EBGRES-2.

5.7.1. Blacklisting and discrete dynamic delay function. Unpredictable and heterogeneous links in a wireless sensor network require techniques to avoid low delivery rate and high delivery cost. Blacklisting bad links has been shown as an efficient mechanism to avoid ‘weak links’ [40]. Blacklisting is a technique that prevents low quality links from being considered for path selection [41]. With blacklisting, all nodes collect statistics about delivery rates with their neighbors. Links with loss rate below a configured threshold are ignored, inbound and outbound packets on those links are dropped. By avoiding tenuous links, blacklisting can improve end-to-end reliability, although ignoring links risks partitioning the network.

For any node \( v \in R_u \), node \( v \) is blacklisted from participating in the contention for acting as packet relay for node \( v \) if \( PRR(u, v) < \eta \).

Let \( B(u) \) be the set of remaining nodes in \( R_u \) after blacklisting. Clearly,

\[
\max_{v \in B(u)} \frac{|vf_u|}{PRR(u, v)} \leq \frac{r_s(u)}{\eta}.
\]

A DDFD function similar to the one in EBGRES is then used to reduce the number of ACK messages broadcasted. The principle of this discrete delay function is the same with that in EBGRES. The nodes in \( B(u) \) are divided into \( n \) sets \( S_1, S_2, ..., S_n \) based on the following rule: If

\[
v \in S_i, \frac{(i-1) \cdot r_s(u)}{\eta} \leq \frac{|vf_u|}{PRR(u, v)} < \frac{i \cdot r_s(u)}{\eta}
\]

The delay for node \( v \) to broadcast its ACK message after receiving a DATA message from \( u \), denoted by \( \delta_{v \rightarrow u} \), is defined as follows:

\[
\delta_{v \rightarrow u} = 2(m - 1) \cdot \gamma \cdot r_s(u) + \gamma \cdot \frac{|vf_u|}{PRR(u, v)}
\]

where \( m = \left\lfloor \frac{n \cdot \eta \cdot |vf_u|}{r_s \cdot PRR(u, v)} \right\rfloor + 1 \)

The delay setting ensures that the neighbors in \( S_1 \) broadcast the ACK messages first. Within each set, the neighbor that has a smaller \( |vf_u|/PRR(u, v) \) is assigned with a shorter delay. Furthermore, the ACK message broadcasted by a node in one set can be snooped by the nodes in other sets before they broadcast their own ACK messages. It is worth pointing out that the number of messages needed to be broadcasted is minimized when there is only one node in the innermost nonempty set.

6. CONCLUSIONS

In the paper we presented in detail an online geographic routing scheme called EBGRES, which can provide fully stateless, energy-efficient source-to-sink, and scalable routing approach that has communication overheads without the need to maintain neighborhood information with perpetual energy supply. EBGRES makes routing decisions locally by jointly taking into account multiple factors such as energy availability, packet advancement to destination, energy available on the node with energy harvesting capability, node positions, and duty cycle of the node. We proved that EBGRES is loop-free in greedy forwarding mode, and established the lower and upper bounds on hop count for source-to-sink routing for WSNs. We established the upper bound on energy consumption for source-to-sink data delivery under EBGRES, assuming no packet loss and failures in greedy forwarding mode when the number of nodes increases. We extended EBGRES to a lossy wireless sensor network to provide energy-efficient source-to-sink routing in the presence of unreliable communication links using packet reception ratio and we introduced a new routing metric. We prove that the extended algorithm EBGRES-2 theoretically performs in a similar way to EBGRES. Availability of perpetual energy has an effect of maintaining the upper and lower bounds for energy consumption, yet it will facilitate an increase in the workload and lifetime of the sensor nodes. Future work involves simulating the algorithm under realistic and practical scenarios. A performance comparison study with other routing algorithms such as BOSS, EBGR, and GPER is of great interest. We will investigate the performance of the algorithm under increased mobility of the nodes, multimedia applications, increased in the number of nodes, and address end-to-end QoS challenges within the network.

ACKNOWLEDGEMENTS

We thank the MIH Media Laboratory at Stellenbosch University, South Africa for funding this research. We also express our gratitude to the anonymous reviewers for their comments which helped us to improve the presentation of this paper. Sherali Zeadally was partly funded by a District of Columbia NASA Space grant.

REFERENCES


