SGF: A state-free gradient-based forwarding protocol for wireless sensor networks.

Article · January 2009

Source: DBLP

Guoliang Xing
Michigan State University

SGF: A State-Free Gradient-Based Forwarding Protocol for Wireless Sensor Networks

PEI HUANG
Michigan State University
HONGYANG CHEN
The University of Tokyo
GUOLIANG XING
Michigan State University
and
YONGDONG TAN
Southwest Jiaotong University

Limitation on available resources is a major challenge in wireless sensor networks. Due to high rates of unexpected node/link failures, robust data delivery through multiple hops also becomes a critical issue. In this article we present a state-free gradient-based forwarding (SGF) protocol to address these challenges. Nodes running SGF do not maintain states of neighbors or network topology and thus can scale to very large networks. Without using routing tables, SGF builds a cost field called *gradient* that provides each node the direction to forward data. The maintenance of gradient is purely driven by data transmissions and hence incurs little overhead. To adapt to transient channel variations and topology changes, the forwarder of a routing node is selected opportunistically among multiple candidate nodes through a distributed contention process. Simulation results show that SGF achieves significant energy savings and outperforms several existing data forwarding protocols in terms of packet delivery ratio and end-to-end delay.

14

Categories and Subject Descriptors: C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless communication; distributed networks*; C.2.2 [**Computer-Communication Networks**]: Network Protocols—*Routing protocols*; C.3 [**Computer Systems Organization**]: Special-Purpose and Application-Based Systems—*Real-time and embedded systems*

Part of this article was presented in *Proceedings of the 5th Annual IEEE Consumer Communcations and Networking Conference (CCNC'08)*[Huang et al. 2008]@IEEE 2008.
Authors' addresses: P. Huang, G. Xing, Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824; email: dekstpei@gmail.com; glxing@msu.edu; H. Chen, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 153-8505 Japan; email: hongyang@mcl.iis.u-tokyo.ac.jp; Y. Tan, School of Electrical Engineering, Southwest Jiaotong University, Chengdu Sichuan, 610031 China; email: ydtan@home.swjtu.edu.cn.

DOI 10.1145/1498915.1498920 http://doi.acm.org/ 10.1145/1498915.1498920

General Terms: Algorithms, Design, Performance, Reliability

Additional Key Words and Phrases: Sensor networks, power control, energy conservation, state-free, gradient, robust performance

ACM Reference Format:

Huang, P., Chen, H., Xing, G., and Tan, Y. 2009. SGF: A state-free gradient-based forwarding protocol for wireless sensor networks. ACM Trans. Sensor Netw., 5, 2, Article 14 (March 2009), 25 pages. DOI = 10.1145/1498915.1498920 http://doi.acm.org/10.1145/1498915.1498920

1. INTRODUCTION

Wireless sensor networks (WSNs) are of increasing interest in many military, civil, and scientific applications such as battlefield surveillance, fire warning, biological habitat monitoring, etc. In these applications, a large number of low-power, low-cost sensor nodes collectively sense the physical environment and forward data samples through multiple hops to the base station.

Designing data forwarding protocols for WSNs faces several major challenges. Operating in an unattended manner, WSNs suffer high rates of node/link failures. Harsh conditions and malicious attacks can easily cause node damage. In addition, low-power wireless links are inherently lossy due to multipath fading and interference. Communication protocols in WSNs must be robust in the face of constantly changing channel conditions and network topologies. Moreover, sensor nodes are often powered by batteries and have limited memory and bandwidth. A data forwarding protocol thus must be energy efficient and incur extremely low overhead.

Numerous approaches have been proposed to achieve robust routing performance in multihop wireless (sensor) networks. Several routing protocols [Couto et al. 2003; Dong et al. 2005; Banerjee and Misra 2002; Li et al. 2006; Woo et al. 2003] yield high data delivery reliability and throughput by incorporating link quality in the routing metric. However, they incur high memory and bandwidth overhead, as nodes must keep track of the quality of links in the neighborhood in order to choose the best forwarders. Multipath routing [Royer and Perkins 1999; Lee and Gerla 2001, 2000; Marina and Das 2001; Ye et al. 2005] is another approach to achieve robust multihop data delivery. However, maintaining multiple routes simultaneously increases control overhead and wastes network energy.

In the aforementioned approaches, nodes must rely on routing tables for data forwarding. The scalability of such *state-based* solutions is often questionable, as the memory and bandwidth overhead incurred by each node increases with the neighborhood size. Routing tables are eliminated in several geographic routing protocols [Blum et al. 2003; Heissenbüttel et al. 2004] which opportunistically select forwarders based on neighbors' locations. A limitation of these protocols is that location information may not be available or accurate in many applications. ExOR [Biswas and Morris 2005] is an *opportunistic* routing protocol without dependence on location information. However, designed to achieve high throughput in wireless mesh networks, it transmits data in batches and requires complex coordination between multiple forwarders. The complexity and high control overhead makes it ill suited for WSNs. Moreover, energy efficiency is not a design objective of ExOR.

This article presents SGF, a state-free gradient-based forwarding protocol designed for WSNs. SGF maintains a cost field called *gradient* that provides each sensor node the direction to forward data toward the sink. SGF does not maintain states of neighbors or network topology and hence is scalable with network size. To conserve energy, the gradient of a node is established based on the minimum energy consumption of transmitting a packet from node to sink. In the data transmission stage, the forwarder of a routing node is selected among multiple candidate nodes through a distributed contention process. The probability that a candidate node wins the contention depends on the node's gradient, channel condition, and remaining energy.

SGF is particularly suitable for large-scale WSNs with dynamic channel conditions and unreliable nodes. With the gradient mechanism, nodes running SGF do not need to constantly maintain routing or neighborhood tables. The gradient updates are purely driven by data transmissions and hence incur little overhead. Choosing forwarders opportunistically enables SGF to be resilient to node/link failures. As a result, the data forwarding routes found by SGF account for both long-term energy efficiency (by the gradient) and transient link conditions (by the forwarder contention mechanism). SGF also includes several components for handling practical issues, such as routing voids in network topology and the severe hidden/exposed terminal problem due to variable transmission power. Our simulation results show that SGF can achieve satisfactory performance in terms of energy consumption, data delivery ratio, and end-to-end delay.

The rest of the article is organized as follows. Section 2 reviews related work. Section 3 describes the channel model used in this work. Section 4 presents the details of gradient establishment. In Section 5, our novel state-free gradient-based forwarding approach is discussed. Section 6 offers simulation results. We finally conclude the article in Section 7.

2. RELATED WORK

In this section, we review the prior research of WSNs that addresses the issues of energy conservation, robustness, and routing.

There exist two basic ways to conserve energy consumption in WSNs. *Power saving* mechanisms allow a node to enter the *sleep state* by powering off its wireless network interface or even the whole system. Power saving mechanisms have been widely used in MAC (Medium Access Control) [Singh and Raghavendra 1998; Ye et al. 2002] and topology control protocols [Chen et al. 2001; Xu et al. 2003, 2001; Ye et al. 2003]. An alternative method of energy conservation is to use *power control* schemes which reduce transmission power of wireless links [Doshi et al. 2002; Gomez et al. 2003; Xing et al. 2005]. In addition to energy savings, power control can potentially lead to higher network throughput by reducing interference and improving spatial channel reuse [Monks et al. 2001; Muqattash and Krunz 2005, 2003]. Power control is also a fundamental technique employed by many topology control algorithms [Santi 2005; Ramanathan and Rosales-Hain 2000; Li et al. 2005]. SGF proposed in this article employs power control to minimize the transmission energy of data routes. Our simulation results show that SGF can easily integrate power saving mechanisms to further reduce network energy consumption.

Several routing protocols employ the *expected transmission count (ETX)* [Couto et al. 2003] between two nodes as the routing metric. It is shown that routes with small ETXs can yield high data delivery throughput [Dong et al. 2005; Banerjee and Misra 2002; Li et al. 2006; Woo et al. 2003; Couto et al. 2003]. However, these protocols need to proactively update routing information (e.g., ETX to the sink), resulting in high bandwidth and memory overhead. Moreover, they are not designed to scale to dense networks as nodes must periodically update information about neighbors (e.g., quality of links).

Multipath routing is another approach to achieve robust multihop data delivery [Royer and Perkins 1999; Lee and Gerla 2001, 2000; Marina and Das 2001; Ye et al. 2005]. However, transient link loss due to multipath fading and interference can considerably affect the efficiency of multipath routing protocols [Jain and Das 2005]. For instance, temporarily bad channel conditions may lead to multiple transmission retries before the network layer switches to an alternate route, which wastes network bandwidth and increases communication delay. Moreover, maintaining multiple paths simultaneously increases control overhead and wastes network energy.

Gradient-routing [Han et al. 2004; Schurgers and Srivastava 2001; Ye et al. 2005; Intanagonwiwat et al. 2003] reduces the overhead of route maintenance by taking the advantage of the "many-to-one" traffic pattern of WSNs. Gradient is a state representing the direction (toward neighboring nodes) through which the sink can be reached [Han et al. 2004]. It can be set up according to different information, such as hop count, energy consumption, or physical distance. Existing gradient-routing protocols like GRAB [Ye et al. 2005] and Directed Diffusion [Intanagonwiwat et al. 2003] heavily rely on periodic flooding to maintain the gradient, resulting in high bandwidth and energy overhead. A key novelty of SGF is that gradient maintenance is purely driven by data transmissions without periodic flooding and hence incurs little overhead. For gradient establishment, SGF extends an existing algorithm [Ye et al. 2001] by addressing some practical issues.

The work most related to this article is *opportunistic routing* that selects forwarders with little knowledge of the existence of neighbors. Opportunistic routing protocols fall into two basic classes: position-based and topology-based protocols. Several routing protocols [Füssler et al. 2003; Blum et al. 2003; Heissenbüttel et al. 2004] confine the choices of forwarders in a geographic zone such as a sector toward the destination node. Neighbors of the routing node compete to be the forwarder based on their locations and possibly other information like remaining energy. However, the dependence on location information limits the applications of these protocols, since location information may not be available or is inaccurate [Shah et al. 2005; Witt and Turau 2006] in many applications. Cao and Abdelzaher [2006] propose to eliminate the dependence on location information by using logical coordinates of nodes. However, the effectiveness of this approach in opportunistic routing has not been studied.

ExOR [Biswas and Morris 2005] is an opportunistic routing protocol without dependence on location information. It assigns different priorities to candidate forwarders according to their ETXs before reaching the destination. Candidates then forward their received packets in order of priority. Both ExOR and SGF are topology-based opportunistic routing protocols without dependence on location information. However, ExOR is not suitable for WSNs for the following reasons. First, ExOR transmits packets in batches in order to achieve high throughput in wireless mesh networks. It heavily utilizes the state information (e.g., the forwarder list and the map of received packets) carried in data packets to achieve robust delivery. However, traffic load in WSNs is low and bursty (event-driven). A batch-based transmission mode would introduce significant buffering delays. Secondly, instead of selecting one forwarder before data transmission, ExOR identifies multiple forwarders based on their reception performance after a data batch is sent. Coordination among forwarders is then needed to avoid duplicate forwarding. Although this strategy is efficient for large data packets, it incurs high control overhead when data packets are small. Therefore, the design of ExOR cannot be directly ported to WSNs.

3. CHANNEL MODEL

The amount of transmission power required for a node i to successfully send a signal to node j depends on the channel *gain* ($G_{i,j}$) between i and j, which models the attenuation of transmission power over distance [Monks et al. 2001]. Assuming that the transmission power of the sender is P_t, the signal power received at the receiver at a distance d from the sender, P_r, can be derived from the large-scale and small-scale propagation models [Rappaport 2001]. In the large-scale propagation model, two path-loss field regions are defined:

—the region where the gain drops with $1/d^2$ (inside the Fresnel zone or within the crossover distance);

$$P_R = \frac{G_t \cdot G_r \cdot \lambda^2}{(4\pi \cdot d)^2 \cdot L} \cdot P_t \tag{1}$$

—the region where the gain is proportional to $1/d^4$ (outside the Fresnel zone or beyond the crossover distance)

$$P_R = \frac{G_t \cdot G_r \cdot {h_t}^2 \cdot {h_r}^2}{d^4 \cdot L} \cdot P_t \tag{2}$$

where P_R denotes the received power computed by the large-scale propagation model. It depends on the transmission power P_t, the gain of the transmitter and receiver antenna (G_t, G_r), the wavelength λ, and the distance d between sender and receiver, as well as a system loss coefficient L when $d < d_{Thresh}$ where $d_{Thresh} = 4\pi \cdot h_t \cdot h_r/\lambda$; otherwise, it is also related to the heights of both antennas over the ground (h_t, h_r).

In the small-scale propagation model, the received power P_r can be expressed as the product of P_R and time-varying factor K with known statistical characteristics: $P_r = K \cdot P_R = G_{i,j} \cdot P_t$ [Jain and Das 2005]. In our protocol design, we measure the actual gain $G_{i,j}$ based on the transmission power and the received power of the control packet: ORTS (see Section 5.1). Due to the existence of time-varying factor K, channel gain is measured right before data transmission. The channel gain measured for the ORTS can also be valid for the following data packet if the coherence time[1] is in the order of tens of milliseconds [Jain and Das 2005] and retransmission is employed to recover a lost packet, even if the channel fade state changes from the point when DATA is transmitted.

The minimum transmission power required to send a data packet from node i to node j is then calculated as follows, where RX_{Thresh} denotes the minimum signal power for successfully decoding a packet.

$$P_t^{min} = \frac{RX_{Thresh}}{G_{i,j}} \tag{3}$$

We note that Eq. (3) does not require any location or distance information.

4. GRADIENT ESTABLISHMENT

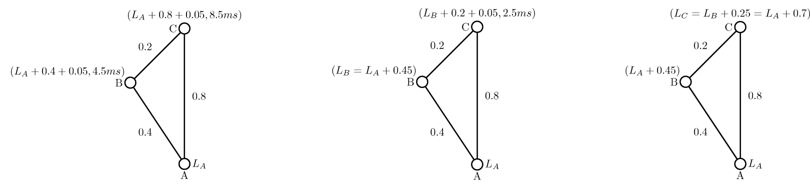
Gradient is a state representing the direction (toward neighboring nodes) through which a destined sink can be reached. It can be set up according to different information, such as hop count, energy consumption, or physical distance. In most existing gradient-based routing protocols [Intanagonwiwat et al. 2003; Schurgers and Srivastava 2001; Ye et al. 2005], the gradient needs to be refreshed by periodic flooding of some forms of control packet, resulting in excessive overhead in terms of energy and bandwidth consumption [Han et al. 2004]. In SGF, energy consumption is used as a metric for establishing *gradient* so that data can be transmitted along the energy-efficient path to the sink. Moreover, gradient needs to be established only once when the network is initialized. The updates of gradient are purely driven by data transmissions without flooding and thus incur little overhead.

4.1 Cost Field

In most traditional routing protocols, paths are computed to minimize hop count or end-to-end delivery delay. When transmission power of nodes is adjustable, the total amount of energy consumed in transmitting a packet along a path is a better routing metric than hop count [Singh et al. 1998]. Several power-aware routing protocols adopt a *redirecting* strategy, where an overhearing node serves as the redirector for a direct one-hop transmission when it finds that an energy reduction can be achieved [Doshi et al. 2002; Gomez et al. 2003]. A drawback of such a strategy is that several iterations are often needed to converge to a stable route between two nodes.

Ye et al. [2001] proposed an efficient backoff-based cost field setup algorithm which can find the optimal costs of all nodes to the sink with the overhead of one message per node. SGF adopts a similar algorithm. In Section 4.3, we describe improvements to the algorithm that address several important practical issues. At each node, the *cost* is defined as the minimum total energy consumption to

[1]The coherence time defines the time duration over which the channel impulse response is essentially invariant [Rappaport 2001].



(a) at time T, after A's advertisement

(b) time T+4.5, after B's advertisement

(c) time T+7, after C's advertisement

Fig. 1. An illustrative example for backoff-based cost field establishment.

send a packet from itself to the sink. The cost includes the energy consumed for both packet transmission and reception. Different from traditional routing protocols, a node does not need to maintain any information of neighbors such as a routing table. Therefore, the memory overhead is minimal.

4.2 Backoff-Based Cost Field Establishment Algorithm

A straightforward way to set up the cost field would be through networkwide flooding. However, flooding will lead to high message overhead and energy consumption. To address this issue, a backoff-based scheme is proposed in Ye et al. [2001] to allow a node to defer its broadcast properly. This scheme incurs much lower overhead than the naive flooding scheme where a node rebroadcasts whenever it sees a cost reduction. We now illustrate the basic idea using the following example.[2]

—At time T, sink A broadcasts an ADV (advertisement) message containing its cost $L_A = 0$ and both node B and C receive the message. As the initial cost of B and C is ∞, B reduces its cost to $L_A + 0.4W + 0.05W$, where 0.4 W is the link cost between A and B and 0.05 W is the reception cost defined in Section 6.2. At the same time, B sets a backoff timer that expires after $\gamma * 0.45W = 4.5ms$. γ is a constant set to $\gamma = 10ms/W$ in Ye et al. [2001]. It is suggested in Ye et al. [2001] that γ should be the same order of magnitude as the transmission, propagation, and processing delays. However, we observed that it should be much greater, as explained later. Similarly, C's cost is set to $L_A + 0.8W + 0.05W$ and a backoff timer is set to expire after $\gamma * 0.85 = 8.5ms$. Compared with flooding, the backoff-based scheme incurs lower message overhead. For instance, both B and C would broadcast immediately in flooding after their costs become smaller than ∞.

—At time T + 4.5 ms, node B's backoff timer expires and it broadcasts an ADV message containing its minimum cost L_B. When node C hears it, C resets its cost to $L_B + 0.25W$ as C's current cost $L_C = L_A + 0.85W > L_B + 0.2W + 0.05W = L_A + 0.7W$. Node C then resets its backoff timer to $\gamma * 0.25W = 2.5ms$ without broadcasting a message. On the contrary, it has to advertise a second message at this time if flooding were to be used.

[2]This example is similar to an example introduced in Ye et al. [2001] and is included here for completeness.

—Finally, at time T + 7 ms, C's timer expires and node C sets its cost to $L_C = L_B + 0.25W = L_A + 0.7W$, and broadcasts an ADV message with this new minimum cost.

The preceding example shows that the backoff-based scheme suppresses nonoptimal advertisement messages. As a result, each node broadcasts only once with the optimal cost. During the propagation of ADV messages, each node calculates its minimum cost to the sink.

4.3 Practical Issues

In this section, we discuss several improvements to the backoff-based cost field establishment algorithm [Ye et al. 2001]. We observed that when the sum of transmission, propagation, and processing delays is substantial, the order of broadcasts along a minimum-cost path may be altered. We use Figure 1 to illustrate this observation. Suppose nodes have a $10ms$ delay and γ may be set to $10ms/W$ according to Ye et al. [2001]. After B broadcasts an ADV, it will take $10ms$ to complete the transmission and processing at node C.

—At time T, node A broadcasts an ADV.
—At time T + 10 ms, node B sets its cost to $L_A + 0.45W$ and sets a backoff timer that expires after $\gamma * 0.45W = 4.5ms$. Similarly, C sets its cost to $L_A + 0.85W$ and sets a backoff timer that expires after $\gamma * 0.85W = 8.5ms$.
—At time T + 14.5 ms, node B's backoff timer expires and it broadcasts an ADV message that contains its minimum cost L_B. It will take $10ms$ to complete the transmission and processing at node C (at time T + 24.5 ms).
—At time T + 18.5 ms, however, the backoff timer of node C has expired and it broadcasts an ADV message before it notices the lower cost from node B.
—At time T + 24.5 ms, node C will complete the processing of the ADV from node B and sets its own cost to $L_B + 0.25W$. It must broadcast again after $2.5ms$.

The previous example shows that the value of $\gamma * min\_diff$ should be larger than the sum of transmission, propagation, and processing delays, where min_diff denotes the minimum cost difference between any two nodes. Since $50mW$ will be required to drive the transceiver circuitry in our energy model (see Section 6.2), the minimum cost difference between any two nodes would be $min\_diff = 0.05W$.

The delay in the MAC layer should also be considered, since MAC introduces exponential backoff. Suppose nodes have a $2ms$ delay and the maximum total delay might be $c * 2ms$, where c is a constant that compensates for the delay caused by failure of channel contention. As a conservative estimation, we set $c = 3$. The maximum total delay would be $6ms$.

Let $\gamma = 200ms/W$, which ensures $\gamma * min\_diff = 10ms > 6ms$. In Figure 1, if $L_B = L_A + 0.4W + 0.05W$ and $L_C = L_A + 0.45W + 0.05W$ when node B and C hear the ADV from node A, they will defer their broadcasts 90ms and 100ms, respectively. After waiting for at most 6 ms following B's broadcast of an ADV (at time $T + 90ms$), node C will complete the reception and processing

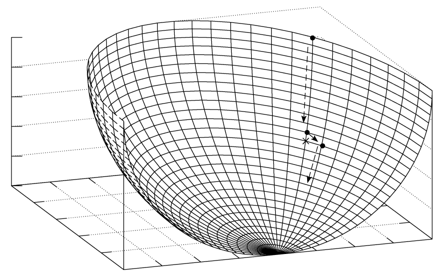


Fig. 2. The "shape" of the cost field.

and compare its current cost with $L_C = L_B + L_{BC} + 0.05W$ before its previously set timer expires (at $T + 100ms > T + 96ms$).

A packet relayed by node B to C will consume an additional power of at least 50mW. Therefore, if a direct transmission will consume less than $L_B + 0.05W$, there is no need to wait for the ADV message from node B. Otherwise, there would be enough time to wait for the ADV message from node B to reach C. This is because even if the link cost L_{BC} is not considered, node C will defer at least 10ms later than node B, which is longer than the sum of transmission, propagation, and processing delays (6 ms).

In the cost field setup algorithm, we assume that the channel gain between two nodes is approximately the same in both directions. This assumption is realistic as long as the multipath effects are small [Monks et al. 2001]. Several recent empirical studies showed that this assumption may not hold in some scenarios [Couto et al. 2003; Son et al. 2004]. However, this issue does not significantly affect the efficiency of our approach. The consequence of asymmetric channel gains is that a node may estimate a suboptimal cost to the sink. However, the actual next-hop node in the data forwarding stage is chosen from a set of neighbors through a contention process (to be discussed in Section 5.1) in which the cost of each node will be updated according to the actual directional gain.

5. STATE-FREE GRADIENT-BASED FORWARDING

5.1 Main Idea of Our Algorithm

Once the gradient is established, messages then can flow along an energy-efficient path to the sink (see Figure 2). The process of message delivery mimics a natural phenomenon that mountain streams flow from the mountaintop down to the bottom of valley [Ye et al. 2001]. When a link on a message route is broken, messages will immediately choose another route to the sink from the node which cannot reach the next hop, which is an analogy of a mountain stream bypassing hillocks on the mountainside. Our goal is to enable a node quickly to switch to the next best forwarder with minimum overhead when a link is broken.

We make several modifications to the Distributed Control Function (DCF) of the IEEE 802.11 standard. In 802.11, there exists a SIFS (Short Interframe Space) between a RTS (Request To Send) and a CTS (Clear To Send). In our

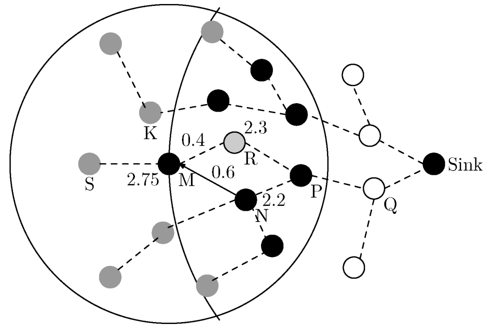


Fig. 3. Data forwarding process of SGF.

proposed protocol, an additional dynamic delay is introduced to allow candidate next-hop nodes to contend to be the forwarder. RTS/CTS frames are replaced by our new *Open RTS* and *Competing CTS* (ORTS/CCTS). Several beaconless geographic routing protocols [Blum et al. 2003; Füssler et al. 2003] also employed timer-based forwarder contention mechanisms. Our design has several key differences. First, SGF takes into account the channel condition (link cost) in the forwarder contention algorithm. Secondly, SGF aims to minimize energy consumption and does not reply on any geographic information.

Figure 3 illustrates our basic idea using an example. Suppose "S-M-R-P-Q-Sink" is the current minimum-cost path from S to the sink when the link from M to R breaks. When node M has packets to send, it chooses a forwarder as follows. First, M broadcasts an *ORTS* which carries the minimum cost of the node to the sink (2.75 W). A neighboring node hearing the ORTS sets a *CCTS_Response* timer if its cost is smaller than that of the sender (e.g., the black nodes within M's transmission range in Figure 3). The timer defines the amount of time that the node must wait before responding to the ORTS. The function used to calculate the CCTS_Response waiting time will be discussed in Section 5.3. If we only consider energy cost, Eq. (4) can be used to calculate the delay of node N.

$$CCTS\_Response = SIFS + Max\_Delay \times \frac{(L_N + C_{M,N}) - L_M}{C_{Max}} \tag{4}$$

—*SIFS* is equal to $10\mu s$ in IEEE 802.11b.

—*Max_Delay* could be $40\mu s$ so that $SIFS + Max\_Delay \leq DIFS$, where DIFS is the DCF interframe space whose value is $50\mu s$ in IEEE 802.11b.

—C_{Max} denotes the maximum cost of a single hop, which is set to be the energy consumed in transmitting (with the maximum transmission power) and receiving a packet.

—L_N, L_M are the cost values of node N and node M, respectively.

—$C_{M,N}$ is the cost from node M to N including transmission and reception.

According to the aforesaid settings, the maximum one-hop delay introduced by the CCTS_Response timer is $40\mu s$, which is smaller than the DIFS. Therefore, the impact on the end-to-end delay is moderate. Assuming $C_{Max} = 1W$, we can obtain node N's time of delay using Eq. (4).

$$\begin{aligned} CCTS\_Response &= 10\mu s + 40\mu s \times \frac{(2.2 + 0.6 + 0.05) - 2.75}{1} \\ &= 14\mu s \end{aligned}$$

If node R hears the ORTS, it will respond before node N, since it has no additional delay except the fixed SIFS ($10\mu s$). This dynamic delay ensures that the current energy-efficient path continues to be used while other candidate forwarders will respond and a new forwarder will be selected if the link to the current forwarder is broken. Therefore, all possible routes from the current routing node to the sink are implicitly maintained. When node N's CCTS_Response timer expires, it responds to node M with a *Competing CTS* (CCTS) and other neighbors sensing it will cancel their timers. The CCTS includes the currently measured channel gain ($G_{M,N}$) and the minimum cost (L_N) of the node. Receiving a valid CCTS, sender M records the channel gain $G_{M,N}$ and subsequent packets (DATA, ACK) are handled according to 802.11 DCF semantics (DATA→ACK). M sends the packets using the minimal transmission power which ensures the correct reception (see Section 3).

5.2 Distributed Cost Field Management

5.2.1 *Update of Cost Information.* In this section we describe the distributed maintenance of cost field. We begin with the discussion on possible causes of cost changes.

The cost of a routing node is equal to the sum of the cost of its current forwarder and the minimum transmission power to reach the forwarder. SGF computes the transmission power of a link based on the channel gain measured from ORTS packets (see Section 3). However, the measured channel gain may be inaccurate due to several types of errors. First, the measurement of channel gain is affected by noise. For instance, recent experimental studies [Lin et al. 2006] showed that the maximum per-hour variation of the measured signal strength of a link is about 3 dB for the CC2420 radio [Crossbow 2003]. Multiple channel gain measurements can be averaged to reduce the error caused by noise. However, this strategy introduces extra control overhead. The second source of error is due to interference from other senders. When the receiver is in the range of an interferer, the measured reception power includes the attenuated power of both sender and interferer, and hence the estimated channel gain is smaller than the actual value. SGF handles this type of error by measuring channel gain using the ORTS packets that reserve channel access and hence reduce (although they cannot eliminate) interference from other senders.

On the other hand, the measurement error of channel gain may not always cause the change of a node's cost. First, most radio platforms only have a limited number of adjustable power levels in practice. For instance, the CC2420 radio has only 8 different transmission power levels from −25 to 0 dBm at a step of about 3 dB [Crossbow 2003]. Therefore, a measurement error within 3 dB does not change the transmission power level of sender. Secondly, a node will update its cost only when the difference between the new

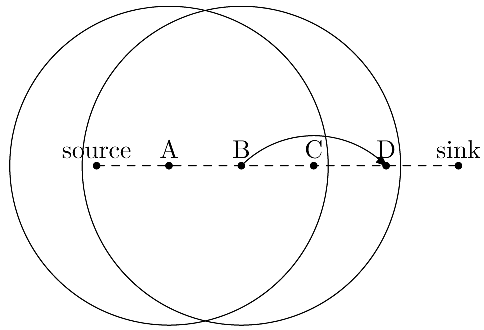


Fig. 4. Illustration of forcible forwarding.

cost and old cost exceeds a threshold, which can tolerate certain measurement variation.

We now discuss the cost update mechanism of SGF when the cost of a node does change, using the example in Figure 3. Suppose node M successfully receives a CCTS from responder N. It extracts N's cost from the CCTS and resets its cost to $L_N + C_{M,N}$, where $C_{M,N}$ is the power cost from node M to N. The cost from a sender to a receiver is the sum of the link cost and a fixed reception cost. If a node's new cost is larger, the current route from this node to the sink may become suboptimal. Consequently, the probability that this node wins the contention among the predecessor's neighbors becomes smaller, which may cause a new forwarder to be chosen for subsequent packets. For instance, when M's cost becomes higher (e.g., due to the quality degradation of channel to R), M will be less likely to relay packets from S (node S might deliver packets through node K the next time).

We can see that the cost increase of a node may make the predecessor automatically choose another forwarder. The cost update is then propagated to upstream nodes in the reverse direction of data forwarding. A key feature of SGF is that each step of cost propagation is purely driven by a data transmission, which is in contrast to many routing protocols that use flooding to update route costs.

5.2.2 *Forcible Forwarding.* In our protocol, packets are expected to flow along the nodes in decreasing order of costs to the sink. However, packets need to be sent from a low-cost node to a high-cost node under certain conditions.

Considering a simple linear topology as illustrated in Figure 4. If the link from node B to node C is broken, node B will forward its packets to node D and update its cost to $L_B = L_D + L_{B,D}$. This may cause $L_B > L_A$ because the long link from B to D has a high cost. Consequently, node B is not eligible to be the forwarder of A, resulting in a packet routing failure, although there exists a route. To deal with this issue, the original cost is recorded in each node at the cost field establishment stage. Those neighbors whose original cost is lower than that of the sender are required to participate in the forwarding competition if they hear a retrying ORTS. Specifically, two costs are maintained by each node: The *update-to-date cost* is used to direct packets to flow along preferable routes,

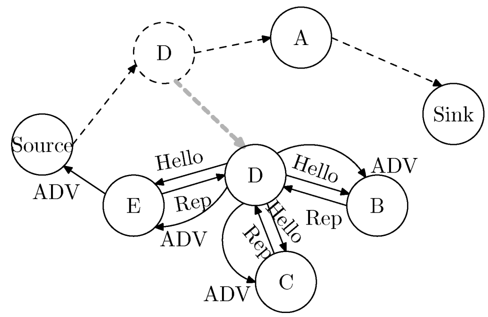


Fig. 5. Gradient reconfiguration.

while the *original cost* is used to make progress forcibly when no preferable routes are found. Note that the cost of A will be updated to $L_A = L_B + L_{A,B}$ (which is larger than the cost of B) when the next packet is forwarded. Therefore, the cost skew of a route will be corrected after a few data transmissions.

5.2.3 *Cost Field Reconfiguration.* From the previous discussion, we can see that the change of the cost field is tracked by data transmissions, and thus periodic *cost field refreshing* is no longer needed. If any node/link fails, the suboptimal node/link will be chosen and the cost information will be updated during the transmission negotiation.

Special attention must be paid to the case of mobile nodes. We here discuss how to handle the semistatic environment [Han et al. 2004] where sensor nodes move intermittently. We also assume that nodes are able to detect their moving state, that is, whether they are on move (moving state) or on pause (pause state). Any node that switches to the moving state is required to reset its cost to ∞ so that it will not be used for data forwarding. After the node stops moving and returns to the pause state, it will try to reconfigure its cost value and broadcast an ADV to notify other nodes of the new cost. We note that SGF handles topology changes caused by node mobility or sleep scheduling in the same way. As SGF does not maintain neighbor information, it cannot distinguish nodes that newly wake up from sleep from those that move to new locations. Similarly, nodes that switch to sleep are also identical to the nodes that move away from their current positions. We now illustrate our basic idea using the example in Figure 5.

When node D completes its movements or wakes up to work, it first broadcasts a hello message to its one-hop neighbors (i.e., nodes B, C, E) and waits for them to respond with their cost values. Neighbors' responses to the hello message will be deferred in proportion to the link cost to avoid collisions. Node D will handle all the responses in the same way it handles ADV messages when calculating its new minimum cost. Then it broadcasts its cost value in an ADV message to update the costs of neighbors.

The channel gain measured by a mobile node may not be accurate (e.g., the measurement is made before a complete stop). Consequently, inaccurate cost information may reduce the energy efficiency of the data forwarding routes around the node. There are two solutions for this problem. First, the node may broadcast several hello messages consecutively and measure the channel gains

Table I. Response Waiting Time

$CCTS\_Response = SIFS + Max\_Delay \times F$
(Note: $Max\_Delay + SIFS \leq DIFS$)
- SIFS: Short Interframe Space (From 802.11)
- DIFS: DCF Interframe Space (From 802.11)
$F = W_C \times C + W_E \times E + W_R \times Random$
- W_C, W_E, W_R: Weight of Cost, Energy, and Random Parameters ($W_C + W_E + W_R = 1$)

using the responses from neighbors. This scheme allows a node to quickly assess the average link cost with several measurements, at the price of extra overhead. Alternatively, the node may overhear packets from neighbors and estimate the average channel gain for a certain amount of time before advertising its cost. This scheme incurs lower overhead while introducing a measurement delay.

We note that the ADV message is usually only propagated within several hops and hence will not cause excessive gradient reconfigurations. For instance, node B will discard the ADV message as its cost is lower than that of node D, and node C will discard it too, as its minimum cost value is derived from node B instead of node D. Therefore, only node E will update its cost and broadcast an ADV message. The aforesaid mechanism allows upstream nodes to quickly learn the new minimum cost and update their costs so that they can utilize this preferable route.

5.3 Setting a Dynamic Response Waiting Timer

It is shown that minimizing the total energy consumption of routes may not lead to the maximum network lifetime [Chang and Tassiulas 2000]. In this section, we discuss how the workload of different nodes is balanced. To this end, the *CCTS_Response* timeout takes into account the energy cost, remaining energy, and an additional random delay (see Table I). Our timeout function extends a similar function, proposed in Blum et al. [2003]. In particular, we adopt energy consumption as the cost metric and do not require any location information.

Random is a real number between 0.0 and 1.0, generated by a uniform distribution function. It is introduced to avoid simultaneous responses and further distribute the workload among forwarding nodes. Cost parameter C denotes the gap between the cost of a node and the minimum cost and hence indicates the level of preference of the node. Energy parameter E is included to consider the remaining energy of nodes. This weight function resembles the timer function proposed in Blum et al. [2003]. In order to avoid collisions among multiple CCTS packets, we later convert it to a logarithmic function.

As discussed in Section 5.1, node N can compute its cost parameter C using Eq. (5) after hearing an ORTS from node M. No additional delay results if node N is the optimal next hop, since $L_M = L_N + C_{M,N}$. Otherwise, the more energy required to relay packets from node N, the later N will respond.

$$C = \frac{(L_N + C_{M,N}) - L_M}{C_{Max}} \tag{5}$$

E measures the fraction of energy that has been used after deployment. From Eq. (6), we can see that a greater value of E will lead to a longer waiting time. This balances the network load and hence nodes across the network consume energy at a roughly even speed.

$$\begin{aligned} E &= \frac{usedEnergy}{initialEnergy} \\ &= 1 - \frac{remainEnergy}{initialEnergy} \end{aligned} \tag{6}$$

W_C, W_E, and W_R are used to tune the weights of different parameters. In our simulations, they are set to 0.5, 0.4, and 0.1, respectively, giving more weight to energy conservation.

We note that other information besides energy consumption can also be included in this weight function. For instance, hop count is used in the gradient-based forwarding in GBR [Schurgers and Srivastava 2001] and Novel GBR [Han et al. 2004]. However, hop count alone is insufficient because there often exist several neighbors with the same hop count to the sink. Consequently, these neighbors will respond to the sender at roughly the same time when contending to be the forwarder. Expected Transmission Count (ETX) would be a better choice than hop count because it accounts for packet retransmissions along the route to the sink. Moreover, the probability that two nodes have the same ETX is small.

5.4 Collision Avoidance

When node density is high, the probability that several neighboring nodes have similar cost and remaining energy cannot be neglected. Although the *Random* parameter introduces a random delay, it only alleviates the problem to a certain extent. One solution is to adopt a topology control algorithm which maintains a constant working node density such as PEAS [Ye et al. 2003]. When no such a topology control algorithm is adopted, we convert the equation in Table I to a logarithmic function that can effectively spread out different nodes' deferral times.

Assuming *minEngergy* Joules have been used to establish the cost field, the minimum value of F would be

$$F\_min = W_E \times (minEngergy / initialEnergy). \tag{7}$$

Let $log_{10}(F)$ denote the logarithm of F, with 10 as the base. Then the equation used to calculate the value of *CCTS_Response* is rewritten as

$$\begin{aligned} CCTS\_Response = SIFS + Max\_Delay \times \\ (log_{10}(F) - log_{10}(F\_min)). \end{aligned} \tag{8}$$

An undesirable property of the logarithmic function is that it cannot well distinguish two large values. However, this has little impact on the efficiency of the forwarder contention mechanism because nodes with large values of F will usually be suppressed by a node with a small value of F.

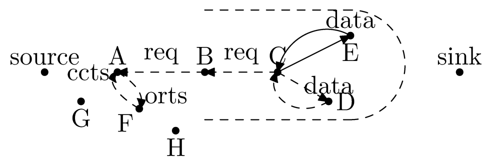


Fig. 6. Retreat scheme.

5.5 Hidden/Exposed Terminal Problem

When wireless nodes transmit at different power levels in DCF of IEEE 802.11, the number of hidden terminals is likely to increase. This is because the sensing/transmission ranges are different for different nodes when multiple power levels are used. Consequently, more collisions and retransmissions can result, due to DCF's contention-based access mechanism, and hence more energy is wasted [Qiao et al. 2003].

We incorporate the *Enhanced Carrier Sensing* (ECS) scheme [Li et al. 2005] to handle the increased collisions caused by power control. ECS was originally proposed to detect the type of erroneous frames and defer nodes' transmissions accordingly. We observed that ECS considerably alleviates the severe hidden/exposed terminal problem caused by variable transmission power. This is because ECS can correctly distinguish different types of frames, even when a node is in the interference range while not in the transmission range of a sender, which is a very common case when different transmission powers are used.

5.6 Recovery from Routing Voids

Routing voids refer to those scenarios where a sender cannot find any forwarder. Routing voids can be easily caused by battery depletion of nodes or malicious attacks. In this section, we discuss a *Retreat* mechanism that can improve data delivery performance in the presence of routing voids. The basic idea is that a packet falling into a routing void should be sent back to the predecessor of the current routing node.

For example, node E in Figure 6 has received a data packet from node C and it then broadcasts an ORTS to find the forwarder. If no response is received when the ORTS timer expires, E will retransmit the ORTS several times. This is because the first failure may also be caused by collisions of responses from neighbors. However, several consecutive failures usually indicate a routing void because the probability of multiple collisions is small. In such a case, the data packet will be sent back to the predecessor (node C) rather than simply dropped. Node E then sets its cost to infinity and goes to sleep. If node C cannot find another forwarder, the packet will be dropped. Node C sets its cost to ∞ and also requests all nodes whose minimum costs are derived from it to explore new forwarders. The process will be repeated until an upstream node (say, A) finds a good forwarder (e.g., F). We note that the retreat mechanism is not designed to ensure the delivery of all data packets. Instead, we aim to inform upstream nodes to switch to other routes before getting into voids. Sleeping nodes will wake up later and reconfigure their costs locally.

6. PERFORMANCE EVALUATION

6.1 Simulation Methodology

This section presents simulation results. We compare SGF against two existing routing protocols: GRAB [Ye et al. 2005] and AODV [Perkins and Royer 1999]. GRAB is a robust data delivery protocol in which several copies of data reports are delivered along a group of *interleaved paths* from source to sink. When a node generates a report, it includes the sum of its own energy cost and a fixed extra energy *credit*. Receiving this packet, a neighboring node further broadcasts it *only if* the sum of its own cost and the link cost is smaller than the total energy budget of the sender and it has not forwarded any copy of the report. The extra energy credit included in data reports will result in multiple paths like an interleaved forwarding mesh between source and sink. The value of *credit* determines the "width" of the forwarding mesh and hence directly affects the robustness and energy consumption of GRAB.

We compare SGF against GRAB for three reasons. First, it is shown in Ye et al. [2005] that GRAB is superior to a widely used sensor network routing protocol, Directed Diffusion [Intanagonwiwat et al. 2003], in terms of both delivery ratio and energy consumption. Secondly, GRAB is a typical multipath routing protocol and the comparison allows us to evaluate how the single-path delivery of SGF would perform when handling node/link failures. Thirdly, the gradient establishment of SGF is an extension to that of GRAB. Moreover, the gradient maintenance of SGF is purely driven by data transmissions while GRAB uses flooding for gradient maintenance. The comparison thus evaluates the effectiveness of our improvements.

As a flooding-based routing protocol, AODV can find short routes with low delays. In other words, AODV trades energy consumption with latency. SGF achieves robust data delivery by sacrificing the end-to-end delay performance (because of the forwarder contention mechanism). Therefore, comparison against AODV allows us to evaluate the impact of our forwarder contention algorithm on end-to-end delivery delay.

We carry out simulations using the *ns-2* network simulator with the CMU wireless extensions [Broch et al. 1998]. Sensor nodes are uniformly distributed in a $1000 \times 500m^2$ field. The initial energy of all nodes is 50 joules. A source node in the left-bottom corner generates an event report every second, with a packet size of 512 bytes, and sends it to the sink node located on the right-top corner. Source and sink nodes have enough energy to remain alive throughout simulations. The credit of GRAB is set to 0.6 to achieve good delivery quality and low energy consumption.

6.2 Energy Model

Many power-aware routing protocols only consider the energy consumption of transmitters while ignoring the energy consumption of receivers. However, it is shown in Heinzelman et al. [2000] and Doshi et al. [2002] that the energy consumption of wireless transmissions depends on both the fixed energy consumption of radio circuitry and the distance between sender and receiver. As a result, a multihop route may consume more energy than the direct hop route if the energy consumption of radio circuitry is substantial.

A simple radio energy model is presented in Heinzelman et al. [2000], where the radio dissipates $E_{elec} = 50nJ/bit$ to drive radio circuitry. It is assumed that energy attenuates with $1/d^2$ within the crossover distance and $1/d^4$ beyond the crossover distance. The transmitting amplifier at the sender consumes $E_{friis\_amp} \cdot d^2$ or $E_{tworay\_amp} \cdot d^4$ in addition to the fixed power to run radio circuitry. $E_{friis\_amp} = 100pJ/bit/m^2$ and $E_{tworay\_amp} = 0.013476pJ/bit/m^4$. Thus, to transmit one bit for distance d, the radio consumes $E_{elec} + E_{amp} \cdot d^{\alpha}$ Jules where $\alpha = 2$ or 4 and to receive the one bit, the radio consumes E_{elec}.

Heinzelman et al. [2000] claimed that the aforementioned parameters are consistent with the current state-of-the-art radio design and thus we adopt them in our simulations. Assuming a bandwidth of $1Mbps$, the power consumption of transceiver circuitry can be calculated as

$$P_{elec} = E_{elec} * 10^6 = 50mW.$$

6.3 Impact of Density Control

We also run a density control protocol called PEAS [Ye et al. 2003] in simulations. PEAS extends the network lifetime by activating only a small number of nodes to maintain the sensing coverage, and redundant nodes are suppressed by working nodes for future use. The primary performance metric is the *distinct event delivery ratio*, which is the ratio of the number of distinct event reports received to the number originally sent. The results are measured every 10 seconds. Figure 7 shows the delivery ratios of GRAB and SGF.

We observed two main drawbacks of GRAB from the simulations. First, it suffers significant data losses due to collisions. The collisions occur mainly when packets from multiple paths arrive at several bottleneck nodes. The situation is especially severe in the vicinity of the sink. The reason is that GRAB only adopts a CSMA (Carrier Sense Multiple Access) mechanism, while no acknowledgment or retransmission is used to ensure reliable delivery between two nodes. It merely relies on multiple copies of data to cope with unexpected node/link failures. However, this strategy leads to significantly more collisions.

Secondly, GRAB also confines nodes' broadcast range in order to save energy. When a node forwards packets, it chooses a broadcast range that covers several nodes with lower costs (by looking up the neighborhood table constructed in the cost field establishment stage). However, after these neighbors deplete their energy, the sender still broadcasts packets using the initially chosen transmission power and thus fails to forward the packets. Furthermore, no mechanism is used to notify the sender of the transmission failure. This causes consecutive event report losses and the energy depletion of a few working nodes. This phenomenon explains several drastic drops of delivery ratio in Figure 7. Although PEAS wakes up sleeping nodes when working nodes die, it takes a certain delay for the sink to notice the big variation of the monitored parameter, namely the success ratio, and then to broadcast an ADV to refresh the cost field.

In SGF, the sender broadcasts the ORTS at the maximum transmission power level, which increases the opportunity of finding a receiver in the desired

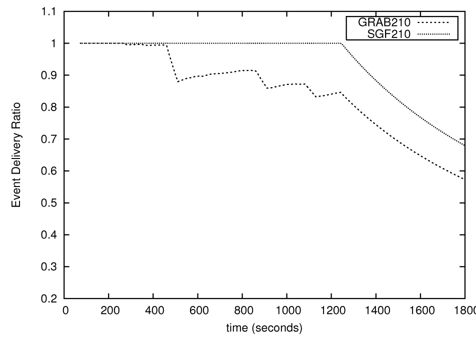
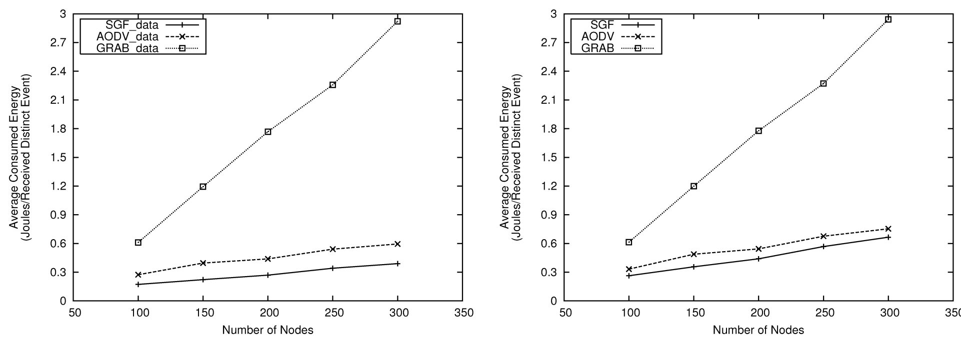


Fig. 7. Success ratio comparison (210 nodes).



(a) Energy consumption of data packets

(b) Energy consumption of both data and control packets

Fig. 8. Energy consumption comparison.

direction. In addition, the actual forwarder is determined dynamically when data is transmitted. Therefore, SGF is resilient to frequent topology changes. Furthermore, in contrast to GRAB, SGF only sends one copy of data from the source, which results in lighter network contention. The integration of ECS (see Section 5.5) also alleviates the hidden/exposed terminal problem. By integrating the aforementioned mechanisms together, SGF yields satisfactory packet delivery performance.

In this set of simulations, the network lifetime is impacted by both PEAS and the evaluated routing protocols. In following simulations, we disable PEAS and solely evaluate the performance of routing protocols.

6.4 Impact of Node Density

We simulated each protocol with five different network topologies, ranging from 100 to 300 nodes. The results are the average of ten runs. Figure 8 shows the average energy consumption per received distinct event as a function of node density.

We can see that GRAB consumes more energy than the single-path protocols AODV and SGF. SGF outperforms AODV, as it employs energy consumption as the routing metric. However, when the topology is sparse, the difference between the multipath and single-path protocols is small and the transmission power control mechanism used by SGF does not yield much energy savings. This is because nodes in SGF must transmit at the maximum transmission power level most of the time in order to keep network connectivity. When node density increases, more nodes are involved in relaying packets in GRAB. As a result, the total energy consumption increases rapidly. On the other hand, nodes in SGF are able to find energy-efficient routes by exploiting the node density. This result demonstrates the advantage of the transmission power control mechanism employed by SGF.

However, the difference of total energy consumption between SGF and AODV does not expand as that of data energy consumption, although SGF consumes 20% to 25% less total energy than AODV, as shown in Figure 8(b). This is for several reasons. First, the overhead of control packets in SGF increases with node density. More and more nodes receive ORTS/CCTS frames, which are larger than the RTS/CTS frames in the original DCF of 802.11 MAC. Secondly, SGF causes more collisions of control packets due to the forwarder contention mechanism. Although the aforementioned mechanisms introduce extra energy consumption, they significantly improve the robustness of SGF in the presence of node/link failures, as shown in following simulation results.

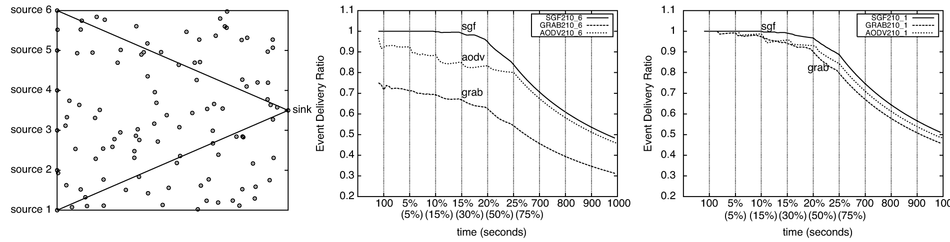
6.5 Impact of Permanent Node Failures

We now study the impact of permanent node failures on SGF. In this set of simulations, six source nodes are placed on the left border and a sink node on the middle of the right border. In such a setup, a failure of most nodes will affect the data delivery performance. Each of the six source nodes generates an event report every 10 seconds and sends them to the sink simultaneously.

We vary the node failure rate from 5% to 25% with a 5% increment every 100 seconds during the nodes' lifetime. In other words, at time 200s, 5% of the nodes (10 nodes) are randomly chosen to shut down. At time 300s, another 10% of the nodes ($210 \times 10\% = 21$ nodes, which is $21/(210-10) = 10.5\%$ of the nodes that are still alive) are turned off (resulting in failures of 15% of the nodes). These nodes are shut down after the reports are sent from sources and before they arrive at the sink.

The data broadcasts in GRAB are staggered in order to reduce collisions. As discussed in Section 5.5, ECS is effective for alleviating the hidden terminal problem caused by variable transmission power. However, we were unable to implement ECS in GRAB because it uses CSMA rather than the 802.11 DCF scheme that exchanges RTS/CTS frames before data transmission. Even though data broadcasts are staggered in GRAB, the collisions are still excessive due to the high traffic load of nodes near the sink (see Figure 9(a)). In AODV, the large number of control packets in the initial route discovery stage results in queue overflow of nodes. Consequently, some data packets are dropped and node failures further decrease the delivery ratio of AODV.

Surprisingly, GRAB delivers fewer event reports than AODV, even though it is designed to achieve high reliability. Each node in GRAB chooses its broadcast



(a) simulation scenario (b) delivery ratio with six source nodes (c) delivery ratio with only one source node

Fig. 9. Delivery ratio comparison with node failures.

range according to the neighborhood table. When only a few nodes fail, the multipath delivery of GRAB ensures that all event reports are received despite collisions. When 10% or more of the nodes fail, the broadcast ranges chosen by nodes become too small, resulting in transmission failures. However, no mechanism is used to notify the sender of the failures. The delivery ratio keeps decreasing until the sink notices the drop and then broadcasts an ADV to refresh the costs of nodes and discover new routes. As a result, the delivery ratio of GRAB is lower than that of AODV in the presence of permanent node failures.

We also simulate the scenario of one source node as in Ye et al. [2005]. Figure 9(c) shows that, although SGF still yields the best performance, its difference with other two protocols is small. This is because the node failure rate is low and the probability that the nodes on the active data route are shut down is low. However, we can see from the results that AODV and GRAB are more sensitive to permanent node failures than SGF.

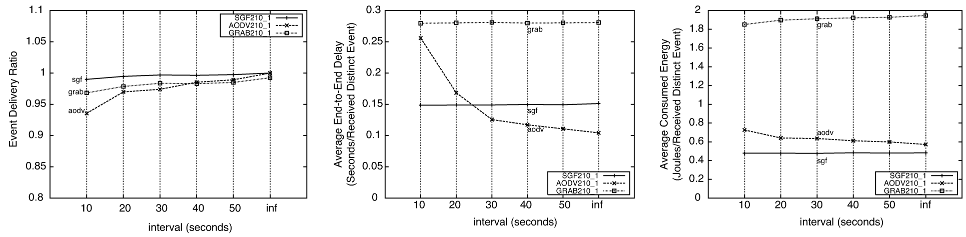
6.6 Impact of Transient Node Failures

We now evaluate the performance of SGF in the presence of transient node failures. Transient node failures also simulate link loss caused by fast fading (multipath effects) or temporary obstacles. The traffic rate is one event per second. Figure 9(c) shows that a 15% node failure rate (17% of the nodes that are still alive) was enough to shut down a considerable fraction of nodes on the active data route. Hence we use a fixed 20% node failure rate in this set of simulations.

The transient node failure model is similar to that in Ye et al. [2005]. Specifically, 20% of the nodes are randomly chosen to turn off for one second, then after certain interval another 20% of the nodes are chosen and turned off. The interval between node failures is varied from 50s to 10s. Under these settings, protocols face lots of dynamics in channel conditions and network topologies.

GRAB is relatively insensitive to data delivery failures and thus its delivery ratio is the lowest in previous simulations. However, transient link loss usually does not result in consecutive data delivery failures. The loss of several data packets does not affect the event delivery ratio of GRAB, since multiple copies of data are transmitted through different routes. On the contrary, transient link loss significantly affects the delivery performance of AODV.

Figure 10(a) shows the event delivery ratio as a function of the interval between node failures. When the interval is infinity (i.e., no nodes fail), the



(a) event delivery ratio (b) average end-to-end delay (c) average energy consumption

Fig. 10. Impact of transient node failures.

event delivery ratio of both SGF and AODV is 100%. However, GRAB does not achieve 100% delivery due to collisions, as discussed earlier. As the node failure interval becomes shorter, the delivery ratio of all three protocols decreases due to more frequent network topology changes. GRAB outperforms AODV due to the existence of multiple data routes. The delivery ratio of SGF is the highest among the three protocols. The reason is that the event delivery of SGF fails only when the sender breaks down and subsequent messages bypass the failing node.

Figure 10(b) shows the average end-to-end delay of the three protocols. We can see that SGF and GRAB are not affected by transient node failures because they dynamically determine forwarders. On the contrary, AODV results in higher delays in the presence of more frequent node failures because the route (re-)discovery process is often triggered to find new routes. When there are no node failures, the average delay of AODV is only about 70% of that of SGF. However, when the topology changes more rapidly, AODV needs to initiate the route (re-)discovery process more frequently. Consequently, the average delay of AODV increases drastically to above 170% of that of SGF. Moreover, the frequent route (re-)discoveries significantly increase the energy consumption. SGF outperforms AODV by 15% to 35% in terms of energy consumption, as shown in Figure 10(c).

Although SGF introduces dynamic delays in the forwarder contention process, the maximum delay is confined within $50\mu s$ (i.e., the duration of DIFS). The high delivery delay of GRAB is due to the contention in channel acquisition and the exponential backoff in collision avoidance. GRAB consumes less energy when the interval between transient node failures is shorter because a smaller number of nodes are involved in data transmissions.

7. CONCLUSION

In this article, we present a state-free gradient-based forwarding (SGF) protocol for robust and energy-efficient data delivery in wireless sensor networks. Without using routing tables, SGF utilizes a cost field called gradient to provide each sensor node the direction to forward sensing data. Gradient is constructed based on the transmission energy consumed before reaching the sink and is refreshed on demand when data is transmitted. The forwarder of a routing node is selected opportunistically among multiple candidate nodes through a distributed contention process. SGF is particularly suitable for large-scale WSNs with dynamic channel conditions and unreliable nodes. Simulation results show that SGF is as robust as a typical multipath routing protocol while achieving more energy savings. Moreover, the on-demand forwarding contention mechanism of SGF has little negative impact on end-to-end communication latency.

In the future, we plan to improve the channel utilization of SGF by further making use of power control, which is critical for achieving high throughput in data-intensive applications. In addition, we will implement and evaluate SGF on a real sensor network testbed.

ACKNOWLEDGMENTS

We would like to thank F. Ye from IBM Research for discussion and providing us the code of GRAB.

REFERENCES

Banerjee, S. and Misra, A. 2002. Minimum energy paths for reliable communication in multi-hop wireless networks. In *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'02)*, 146–156.

Biswas, S. and Morris, R. 2005. ExOR: Opportunistic multi-hop routing for wireless networks. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, 133–144.

Blum, B., He, T., Son, S., and Stankovic, J. 2003. IGF: A state-free robust communication protocol for wireless sensor networks. Tech. rep. CS-2003-11, Department of Computer Science, University of Virginia.

Broch, J., Maltz, D. A., Johnson, D. B., Hu, Y.-C., and Jetcheva, J. 1998. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proceedings of the 4th Annual International Conference on Mobile Computing and Networking (MobiCom'98)*, 85–97.

Cao, Q. and Abdelzaher, T. 2006. Scalable logical coordinates framework for routing in wireless sensor networks. *ACM Trans. Sensor Netw. 2,* 4, 557–593.

Chang, J.-H. and Tassiulas, L. 2000. Energy conserving routing in wireless ad-hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (InfoCom)*. vol. 1. 22–31.

Chen, B., Jamieson, K., Balakrishnan, H., and Morris, R. 2001. Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. In *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom'01)*, 85–96.

Couto, D. S. J. D., Aguayo, D., Bicket, J., and Morris, R. 2003. A high-throughput path metric for multi-hop wireless routing. In *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom'03)*, 134–146.

Crossbow. 2003. Mica/mica2/micaz wireless measurement system datasheets. http://inf.tu-dresen.de/dargie/wsn/slides/students/MICA.ppt.

Dong, Q., Banerjee, S., Adler, M., and Misra, A. 2005. Minimum energy reliable paths using unreliable wireless links. In *Proceedings of the 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'05)*, 449–459.

Doshi, S., Bhandare, S., and Brown, T. X. 2002. An on-demand minimum energy routing protocol for a wireless ad hoc network. *ACM SIGMOBILE Mobile Comput. Commun. Rev. 6,* 3, 50–66.

Füssler, H., Widmer, J., Mauve, M., and Hartenstein, H. 2003. A novel forwarding paradigm for position-based routing (with implicit addressing). In *Proceedings of IEEE 18th Annual Workshop on Computer Communications (CCW'03)*, 194–200.

Gomez, J., Campbell, A. T., Naghshineh, M., and Bisdikian, C. 2003. PARO: Supporting dynamic power controlled routing in wireless ad hoc networks. *ACM/Kluwer J. Wireless Netw. 9,* 5, 443–460.

HAN, K.-H., KO, Y.-B., AND KIM, J.-H. 2004. A novel gradient approach for efficient data dissemination in wireless sensor networks. In *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC Fall)*, vol. 4, 2979–2983.

HEINZELMAN, W. R., CHANDRAKASAN, A., AND BALAKRISHNAN, H. 2000. Energy-Efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Hawaii Inernational Conference on System Sciences (HICSS'00)*, vol. 2, 3005–3014.

HEISSENBÜTTEL, M., BRAUN, T., BERNOULLI, T., AND WÄLCHLI, M. 2004. BLR: Beacon-Less routing algorithm for mobile ad-hoc networks. *Elsevier's Comput. Commun. J. 27,* 11, 1076–1086.

HUANG, P., YANG, X., AND TAN, Y. 2008. A Robust and Energy-Efficient Approach for Image/Video Dissemination in WSNs. In *Proceedings of the 5th Annual IEEE Consumer Communications and Networking Conference (CCNC'08)*.

INTANAGONWIWAT, C., GOVINDAN, R., ESTRIN, D., HEIDEMANN, J., AND SILVA, F. 2003. Directed diffusion for wireless sensor networking. *IEEE/ACM Trans. Netw. 11,* 1, 2–16.

JAIN, S. AND DAS, S. R. 2005. Exploiting path diversity in the link layer in wireless ad hoc networks. In *Proceedings of the 6th IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM'05)*, 22–30.

LEE, S.-J. AND GERLA, M. 2000. AODV-BR: Backup routing in ad hoc networks. In *Proceedings of the IEEE Wireless Communication and Networking Conference (WCNC'00)*, vol. 3, 1311–1316.

LEE, S.-J. AND GERLA, M. 2001. Split multipath routing with maximally disjoint paths in ad hoc networks. In *Proceedings of the IEEE International Conference on Communications (ICC'01)*, vol. 10, 3201–3205.

LI, L., HALPERN, J. Y., BAHL, P., WANG, Y.-M., AND WATTENHOFER, R. 2005. A cone-based distributed topology-control algorithm for wireless multi-hop networks. *IEEE/ACM Trans. Netw. 13,* 1, 147–159.

LI, X.-Y., CHEN, H., SHU, Y., CHU, X., AND WU, Y.-W. 2006. Energy efficient routing with unreliable links in wireless networks. In *Proceedings of the IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS'06)*, 160–169.

LI, Z., NANDI, S., AND GUPTA, A. K. 2005. ECS: An enhanced carrier sensing mechanism for wireless ad hoc networks. *Elsevier's Comput. Commun. 28,* 17, 1970–1984.

LIN, S., ZHANG, J., ZHOU, G., GU, L., HE, T., AND STANKOVIC, J. A. 2006. ATPC: Adaptive transmission power control for wireless sensor networks. In *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys'06)*, 223–236.

MARINA, M. K. AND DAS, S. R. 2001. On-Demand multipath distance vector routing in ad hoc networks. In *Proceedings of the 9th International Conference on Network Protocols (ICNP 01)*, 14–23.

MONKS, J. P., BHARGHAVAN, V., AND HWU, W.-M. W. 2001. A power controlled multiple access protocol for wireless packet networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (InfoCom)*, vol. 1, 219–228.

MUQATTASH, A. AND KRUNZ, M. 2003. Power controlled dual channel (PCDC) medium access protocol for wireless ad hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (InfoCom)*, vol. 1, 470–480.

MUQATTASH, A. AND KRUNZ, M. 2005. POWMAC: A single-channel power-control protocol for throughput enhancement in wireless ad hoc networks. *IEEE J. Selected Areas Commun. 23,* 5, 1067–1084.

PERKINS, C. E. AND ROYER, E. M. 1999. Ad-Hoc on-demand distance vector routing. In *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'99)*, 90–100.

QIAO, D., CHOI, S., JAIN, A., AND SHIN, K. G. 2003. Adaptive transmit power control in IEEE 802.11a wireless LANs. In *Proceedings of the 57th IEEE Semiannual Vehicular Technology Conference (VTC Spring)*, vol. 1, 433–437.

RAMANATHAN, R. AND ROSALES-HAIN, R. 2000. Topology control of multihop wireless networks using transmit power adjustment. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (InfoCom)*, vol. 2, 404–413.

RAPPAPORT, T. S. 2001. *Wireless Communication: Principles and Practice*, 2nd ed. Prentice Hall.

ROYER, E. M. AND PERKINS, C. E. 1999. Multicast operation of the ad-hoc on-demand distance vector routing protocol. In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'99)*, 207–218.

Santi, P. 2005. Topology control in wireless ad hoc and sensor networks. *ACM Comput. Surv. 37,* 2, 164–194.

Schurgers, C. and Srivastava, M. B. 2001. Energy efficient routing in wireless sensor networks. In *Proceedings of the IEEE Military Communications Conference (MilCom)*, vol. 1, 357–361.

Shah, R., Wolisz, A., and Rabaey, J. 2005. On the performance of geographical routing in the presence of localization errors. In *Proceedings of the IEEE International Conference on Communications (ICC'05)*, 2979–2985.

Singh, S. and Raghavendra, C. S. 1998. PAMAS - Power aware multi-access protocol with signalling for ad hoc networks. *ACM Comput. Commun. Rev. 28,* 3, 5–26.

Singh, S., Woo, M., and Raghavendra, C. S. 1998. Power-Aware routing in mobile ad hoc networks. In *Proceedings of the ACM Annual International Conference on Mobile Computing and Networking (MobiCom'98)*. 181–190.

Son, D., Krishnamachari, B., and Heidemann, J. 2004. Experimental study of the effects of transmission power control and blacklisting in wireless sensor networks. In *Proceedings of the 1st Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (IEEE SECON'04)*, 289–298.

Witt, M. and Turau, V. 2006. The impact of location errors on geographic routing in sensor networks. In *Proceedings of the International Conference on Wireless and Mobile Communications (ICWMC'06)*, 76–76.

Woo, A., Tong, T., and Culler, D. 2003. Taming the underlying challenges of reliable multihop routing in sensor networks. In *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys'03)*. 14–27.

Xing, G., Lu, C., Zhang, Y., Huang, Q., and Pless, R. 2005. Minimum power configuration in wireless sensor networks. In *Proceedings of the 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'05)*, 390–401.

Xu, Y., Bien, S., Mori, Y., Heidemann, J., and Estrin, D. 2003. Topology control protocols to conserve energy in wireless ad hoc networks. Tech. rep. 6, University of California, Los Angeles, Center for Embedded Networked Computing. January.

Xu, Y., Heidemann, J., and Estrin, D. 2001. Geography-Informed energy conservation for ad hoc routing. In *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom'01)*, 70–84.

Ye, F., Chen, A., Lu, S., and Zhang, L. 2001. A scalable solution to minimum cost forwarding in large sensor networks. In *Proceedings of the 10th International Conference on Computer Communications and Networks (ICCCN'01)*, 304–309.

Ye, F., Zhong, G., Cheng, J., Lu, S., and Zhang, L. 2003. PEAS: A robust energy conserving protocol for long-lived sensor networks. In *Proceedings of the 23rd International Conference on Distributed Computing Systems (ICDCS'03)*, 28–37.

Ye, F., Zhong, G., Lu, S., and Zhang, L. 2005. GRAdient Broadcast: A robust data delivery protocol for large scale sensor networks. *ACM Wireless Netw. 11,* 3, 285–298.

Ye, W., Heidemann, J., and Estrin, D. 2002. An energy-efficient MAC protocol for wireless sensor networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (InfoCom)*, vol. 3, 1567–1576.

Received August 2007; revised February 2008; accepted May 2008