Interaction of Retransmission, Blacklisting, and Routing Metrics for Reliability in Sensor Network Routing

Omprakash Gnawali[§], Mark Yarvis[‡], John Heidemann[†], and Ramesh Govindan[§]

[§] Department of Computer Science, University of Southern California
941 W. 37th Place, Los Angeles, CA 90089
[†] Information Sciences Institute, University of Southern California
4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90292
[‡] Intel Research & Development, 2111 N.E. 25th Avenue, Hillsboro, OR 97124

gnawali@usc.edu, mark.d.yarvis@intel.com, johnh@isi.edu, ramesh@usc.edu

Abstract— **Unpredictable and heterogeneous links in a wireless sensor network require techniques to avoid low delivery rate and high delivery cost. Three commonly used techniques to help discover high quality paths include (1) link-layer retransmission, (2) blacklisting bad links, and (3) end-to-end routing metrics. Using simulation and testbed experiments, we present the first systematic exploration of the tradeoffs of combinations of these approaches, quantifying the effects of each of these three techniques. We identify several key results: One is that per-hop retransmissions (ARQ) is a necessary addition to any other mechanism if reliable data delivery is a goal. Additional interactions between the services are more subtle. First, in a multi-hop network, either blacklisting or reliability metrics like ETX can provide consistent high-reliability paths when added to ARQ. Second, at higher deployment densities, blacklisting has a lower routing overhead than ETX. But at lower densities, blacklisting becomes less stable as the network partitions. These results are consistent across both simulation and testbed experiments. We conclude that ETX with retransmissions is the best choice in general, but that blacklisting may be worth considering at higher densities, either with or without ETX.**

I. Introduction

Radio links in wireless sensor networks exhibit widely varying reliability over time, space, and from node to node. The radio used in current research platforms have shown widely varying performance over time and space and use very simple CSMA MAC protocols [21], [18], [19]. The drive to minimize node cost and size motivate a minimal hardware and software structure, yet the sensor network as a whole must provide a reliable environment for communications. Recent research has explored several techniques to improve reliability: link-level retransmission (ARQ); blacklisting, *i.e.*, rejecting bad links; and routing using a metric that reflects path reliability. To our knowledge, no current research has carefully explored the *interactions* between these techniques as each strives to improve reliability in different ways. That is the goal of this paper.

Per-hop retransmission (often called ARQ at the MAC layer) is a widely used technique to improve reliability of a given link [2]. Retransmissions are attempted one or more times up to some limit before the packet is declared lost. Using link-level ARQ, losses can be quickly detected and corrected, and even a few per-link retransmissions can greatly improve end-to-end reliability.

Blacklisting is a technique that prevents low quality links from being considered for path selection [7]. With blacklisting, all nodes collect statistics about delivery rates with their neighbors. These delivery rates are used to estimate the quality of all wireless "links". Links with loss rate below a configured blacklisting *threshold* are ignored—inbound and outbound packets on those link are dropped. By avoiding tenuous links, blacklisting can improve end-to-end reliability, although ignoring links risks partitioning the network.

Recent research has proposed the use of link reliability as a metric for routing path selection [4]. Such a metric allows the routing protocol to consider cumulative link reliability over paths, and find the most reliable end-to-

[*] Other names and brands may be claimed as the property of others.

[†] John Heidemann is also partially supported through the NSF Division of Civil and Mechanical Systems, grant number E01-CMS-0112665, and grant number ANI-0220026, "MAC Protocols Specific for Sensor Networks"

[§] Ramesh Govindan's work is supported in part by the NSF Center for Embedded Systems (CCR-0121778). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

end path. Several metrics have been proposed to represent reliability, and we review them in Section III-D. Metric-based routing can incur higher control message overhead if link reliability changes frequently.

These three mechanisms are not mutually exclusive; each approaches the problem of improving end-to-end reliability in a different way. However, to our knowledge, there is no literature that systematically compares these techniques across a range of parameters, both individually and in combination. In this paper, we conduct such a systematic study. This study is complicated by the fact that the parameter space is rather large—each technique can be used with different settings and thresholds. We use simulation to explore the space thoroughly, then validate selected simulation results through testbed experiments.

Our goal is to understand how different techniques can combine to provide "reasonably" reliable end-to-end delivery in the face of lossy links. We assume that applications can tolerate or recover from occasional loss [16], [15], and that the primary source of loss is due to noise and environmental effects, not congestion [14], [17]. These characteristics are typical of many current sensor networks.

The main contribution of this paper is this systematic study comparing the techniques of per-hop retransmission, blacklisting, and metric-based routing and studying their interactions. We identify several key results: One is that per-hop retransmissions is a necessary addition to any other mechanism if reliable data delivery is a goal. Additional interactions between the services are more subtle. First, in a multi-hop network, either blacklisting or reliability metrics like ETX can provide consistent high-reliability paths when added to ARQ. Second, at higher deployment densities, blacklisting has a lower routing overhead than ETX. But at lower densities, blacklisting becomes less stable as the network partitions. These results are consistent across both simulation and testbed experiments. Finally, we have conflicting results about the effects of combining all three mechanisms. Testbed results suggest that moderate blacklisting can reduce the cost of route discovery when added to metric-based routing, however this observation is not supported in simulation. We conclude that ETX with retransmissions is the best choice in general, but that blacklisting may be worth considering at higher densities, either with or without ETX. We describe these results in the context of sensor networks; they are also applicable to multi-hop ad hoc networks.

II. Related Work

There has been a great deal of recent work exploring radio link characteristics in sensor networks and ways to improve reliability. Here we review related work about link quality and three approaches to improve performance: per-hop retransmission, blacklisting, and reliability metrics for path selection.

a) Link evaluation: Two recent papers evaluate radio propagation with sensor-network style radios, Woo and Culler [18] and Zhao and Govindan [21]. Zhao and Govindan observed in their experiments that over 10% links are asymmetric and a third of links have loss rate greater than 30% [21]. They find that these results persist in several environments. Woo and Culler [18] find similar results and investigate remediation strategies (described below).

There is growing evidence that poor link quality can cause problems in multi-hop communication. De Couto *et al.* studied the shortest path algorithm in a network with lossy links and found that this algorithm often selects a path with poor reliability [5].

b) Per-hop Retransmission: Per-hop retransmission is probably the oldest known technique to increase delivery rate on a link that is selected for data delivery [2]. MAC level retransmission (or Automatic Repeat Request, ARQ) is used in the 802.11 MAC to improve delivery rate. ARQ and Forward Error Correcting codes have been proposed to improve per-hop reliability [12]. The goal of per-hop retransmission is to improve the quality of a given link, thus improving whatever path has been selected.

c) Blacklisting: Blacklisting eliminates unreliable, lossy, or asymmetric links from the set of links used for communication. Lundgren *et al.* identified gray zones and suggested that links in this zone should be ignored (blacklisted) while making routing decisions [13]. Ignoring fading links [3], only using links with good signal strength [7], and using power at which a message is received to identify good links and using only good links for routing are different ways in which researchers have implemented blacklisting. Ultimately the goal of blacklisting is to avoid poor-quality links, thus forcing selection of reasonable paths.

d) Reliability Metrics in Routing: All routing protocols use some routing metric to select paths. If the routing metric is selected to represent end-to-end reliability, the routing protocol can identify paths with high reliability. De Couto *et al.* proposed an ETX (Expected number of transmissions) metric that considers forward and backward reliabilities to identify high throughput paths in a network [4]. This work focuses on maximizing throughput in 802.11b-based networks. Yarvis *et al.* proposed using the minimum of forward and backward reliabilities as link metric and using that to find the most reliable path but they note that this results in longer paths [19]. Awerbuch *et al.* proposed minimizing the amount of time a packet uses the network (medium time metric) in a multi-rate radio environment to maximize throughput [1]. Draves *et al.* proposed Per-hop Round Trip Time and Per-hop Packet Pair Delay link-quality metrics but conclude that these metrics perform worse than the ETX metric [6].

e) Interactions of these approaches: We know of only one work [18] that considers how retransmission, blacklisting, and reliability metric help improve data delivery performance. Their study examines the effect of blacklisting (with only two thresholds) on shortest path routing. Furthermore, their study does not examine packet delivery in the absence of per-hop retransmissions. Generally speaking, our work more systematically explores the parameter space by comparing different combinations of these three techniques and quantifying the effect of each technique in the combination. For example, we explore the impact of blacklisting (with five different thresholds) on ETX and the ML metric as well. We also develop a deeper understanding of the reliability metric by studying it at different resolutions. Finally, we control the number of retransmissions as a parameter to our routing protocol; this provides an additional insight on how to achieve a desired delivery rate and delivery cost.

III. Detailed Approaches to Improve Path Reliability

We consider the three approaches to improve path reliability: per-hop retransmission, blacklisting, and reliability-based metrics in routing. This section briefly reviews the specific algorithms we use in our simulations and testbed experiments. Since blacklisting and reliability metrics depend on estimates of link reliability we begin by summarizing how link statistics can be collected.

A. Measuring link reliability

Blacklisting and reliability metrics must estimate link quality. Link delivery rate changes over time due to environment or transient traffic characteristics. Link statistics needs to be reasonably responsive to these changes. Woo and Culler evaluated a range of options for link estimator and neighborhood table management [18]. They identify Window Mean with Exponentially Weighted Moving Average (WMEWMA) to be a good estimator of link quality in a wireless sensor network. One can use active or passive techniques to collect link statistics. Active techniques rely on periodic broadcasts containing link statistics about each neighbor. Passive probing involves piggybacking link statistics to the outgoing data packets. Combinations of active and passive probing are also feasible. We use active probing and WMEWMA estimator in our testbed experiments. Choice of a single probing technique should not favor any given protocol since it affects all reliability techniques equally.

B. Per-hop retransmission

Retransmission is a well known technique to improve the quality of unreliable links. Retransmission is often done at the MAC layer, or it can be done at higher layers, both to the same effect. We vary the number of allowed data retransmissions from zero to three.

In contention-based MACs there is often a higher possibility of collision during the contention period. In our testbed experiments we use S-MAC [20] which includes an RTS/CTS protocol. It is important to distinguish retransmissions of the contention signal from retransmissions of the data. We always allow up to seven attempts at retransmitting RTS signals independent of the number of allowed data retransmissions.

C. Blacklisting

Blacklisting removes unreliable links from the set of links routing layer can use to form a path. Only the links with reliability higher than a *blacklisting threshold* are made available for sending and receiving messages. Blacklisting is usually applied above the link layer and before the message gets to the network/routing layer. Our blacklisting implementation drops incoming and outgoing packets on each link that it determines to have reliability below the specified blacklisting threshold. Note that this approach effectively eliminates asymmetric links from consideration. In this paper, we focus on using blacklisting to reduce the impact of quality variations across different links. Adding hysteresis to blacklisting reduces the impact of temporal variation in link quality increasing the stability of network performance. Exploration of temporal effects is an area for future work. In our experiments, we use a single threshold which classifies a link as a “bad” link as soon as and for as long as its reliability falls below the threshold.

Setting a very high blacklisting threshold results in only highly reliable links participating in route selection, which ultimately helps select a path with high end-to-end reliability. However, a high threshold can also make nodes unreachable if removal of links with lower reliability creates a network partition. Setting the threshold too low allows mediocre links to be selected, which could result in a poor path. In our study, we explore the impact of threshold selection on performance. However, since less-reliable links tend to form less-desirable paths, we examined high and moderate thresholds in greater detail than low thresholds.

D. Reliability Metrics

Path reliability, when used as an end-to-end routing metric, can identify the end-to-end most reliable path between two nodes. By default, many sensor [11] and ad hoc routing protocols use latency or hop count as a metric. Because they do not differentiate paths based on reliability, they tend to select paths with low reliability [5]. To determine path quality, we first quantify the reliability of each link in terms of a metric: the success rate, expected number of retransmission, or signal strength. A routing protocol can then combine these link qualities additively or multiplicatively (depending upon the metric), to select paths with the "best" end-to-end reliability.

A given metric has an associated resolution that limits path differentiation. The resolution is applied when links are measured. For instance, for a success rate metric, a resolution of 20% categorizes all links into five classes with reliabilities 0–20%, 21–40%, 41–60%, 61–80%, and 81–100%. A low resolution metric may reduce the quality of the resulting path by treating an 81% link as equivalant to a 99% link. However at high resolutions (say 1%), routing algorithms can over optimize to accomplish limited improvement, switching from a 97% link to a 98% and a 99% link. Since link qualities are experimentally observed and approximate to begin with, these changes incur the cost of propagating new routes while providing little or no actual change in quality.

We use a variant of ETX [4], also proposed as MT [18] by Woo et al., as a reliability metric. ETX is defined as the expected number of transmissions (including retransmission) for a successful end-to-end data forwarding and hop-by-hop acknowledgment. The following expression shows how to compute the ETX metric for a path p consisting of links a..z with forward reliability of $forward_a$ and backward reliability of $backward_a$ for link a:

$$etx(a) = \frac{1}{forward_a * backward_a}$$

$$ETX(p) = etx(a) + ... + etx(z)$$

Our version of ETX rounds the ETX value for each link to its nearest integer, effectively reducing the resolution of the ETX metric. For example, forward and reverse reliabilities in the range [0.82,1] result in an ETX value of 1, which makes links different in reliability by as much as 0.18 appear identical. With poorer link reliability, ETX becomes more sensitive to small difference in link reliability enabling it to compare links at a higher resolution. Thus the resolution of this reliability metric, while variable, is at most 0.18. This implementation was intended to approximate previously reported results as close as possible.

We also consider end-to-end success rate (SR) as a second reliability metric. We use the SR metric to evaluate the effect of metric resolution on performance because it provides consistent resolution across the range of values. To compute end-to-end success rate we use the product of forward and backward reliabilities of all links in a path as our metric. This metric is similar to the metric proposed in [19], but by taking the minimum of forward and backward reliability, that metric tends to under-estimate link reliability when links are asymmetric. Note a variation that includes only forward reliability is a reasonable alternative when acknowledgements are not enabled (but this variation is not evaluated in this paper).

IV. Evaluation Metrics

To compare protocol alternatives we consider the following metrics:

a) Routing Overhead: Routing overhead provides an estimate of energy cost for finding a path for data forwarding. We compute it by counting packets sent during path discovery. This estimate assumes an energy-conserving MAC protocol is in use so that idle listening does not dominate energy consumption. (An alternative is to count packets received; we do not do that because it is more sensitive to density.)

Link quality estimation involves a periodic exchange of bi-directional link quality estimate with each neighbor and can be an additional source of overhead. We do not measure this cost in our experiments and therefore slightly overestimate the relative cost of ML with retransmissions but no blacklisting. However, for this configuration, data delivery cost is much higher than alternative schemes, our overall results do not change.

b) Path Reliability: Path reliability measures the ratio of successfully delivered messages at the sink to the number sent by sources. Although a high path reliability is desirable, a slightly lower reliability may be tolerable if accompanied by much lower overhead.

c) Path Length: We measure path length in hops from source to sink. Longer path lengths correspond to higher delivery latency. This relationship is approximate, however, since we do not explicitly model MAC-level retransmission costs on latency or energy.

d) Data Dissemination Overhead: This metric captures the cost to send data, and includes retransmissions but excludes routing overhead. This metric is computed by normalizing the total number of data transmissions by the number of successfully delivered messages to reflect the cost of packets that are sent but lost. Assuming an energy conserving MAC, data overhead approximates the energy consumed to send data in the system.

We decided to evaluate routing and data dissemination overheads separately so that our result can be used

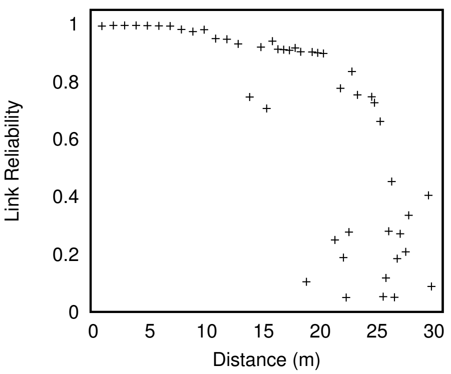


Fig. 1: Reliability vs. distance profile used in the simulation.

to estimate aggregate overheads for applications with different route update and data rates. We also note that retransmissions affect path reliability and data dissemination overhead while blacklisting and reliability metrics impact all evaluation metrics.

V. Simulations

We conducted a simulation study of reliability techniques to systematically explore the parameter space of each mechanism and combinations of the mechanisms. This section reports this exploration; Section VI validates key results in a testbed.

A. Simulation Methodology

We considered the interaction between three techniques: retransmissions, blacklisting, and metric-based routing. We evaluated all 96 combinations of these parameters: 0, 1, 2 or 3 retransmissions; 0%[1], 40%, 60%, 70%, 90%, and 95% blacklisting thresholds; and minimum latency (ML), Success Rate (SR) metric at 1% and 10% resolutions (SR01 and SR10), and expected transmissions (ETX) as routing metrics. For brevity, we summarize the parameters using a three-tuple notation (A, B%, C), where A is the number of retransmissions, B is the blacklisting threshold, and C is the routing metric.

We evaluate these techniques using the one-phase-pull (OPP) variant of Directed Diffusion [9]. Directed Diffusion is a data-centric mechanism for naming, aggregation, and dissemination of information in a sensor network [11]. We chose it because it is used in several sensor network deployments, is freely available, and allows us to observe a specific real protocol. In OPP, the querying node, also called the *sink*, broadcasts a query, also called an *interest*, into the network. Data generated by the *source* nodes are directed back to the sink using previous hop pointers for given query attributes, also called a *gradient*. Queries are re-injected into the network every interest epoch. Intermediate nodes pick up the first interest message they receive and ignore the rest of the interest messages that arrive in the same epoch. By default, nodes with diffusion select a minimum-latency path like the ML metric. To simulate other reliability metrics we extended OPP to encode the routing metric as an additional attribute in the interest message. Nodes update the metric value when they forward the interest message, rebroadcasting interests if the metric improves within a single epoch. While we use this specific protocol, we expect them to be applicable to other ad hoc routing protocols as well.

We conduct our simulations using diffusion release 3.2.0 as a process-level simulator. Packets are sent between nodes as UDP packets. Between each pair of nodes all packets are subject to probabilistic loss as a function of distance based on propagation profiles (Figure 1) from Zhao *et al.* [21].

We consider a 125-node network with nodes placed in a $124 \times 124\text{m}^2$ area using a placement strategy similar to that in [9]. A sink is placed in the lower left sixteenth of the sensor field. We use ten clustered sources; a first source is chosen in the upper right sixteenth of the area, additional sources are taken as the nearest nodes to that source. This constrained source and sink placement allows us to maintain consistent average path lengths from source to sink across different randomly generated instances of the sensor network. To vary node density we changed the number of nodes by placing 45, 65, and 125 nodes in the given area. Assuming a nominal radio range of 30m and uniform node density, these placements result in topologies with an average of 8, 12, and 23 neighbors per node. For a given density, we generated 20 topologies with random node placements and simulated all 96 parameter combinations on those topologies.

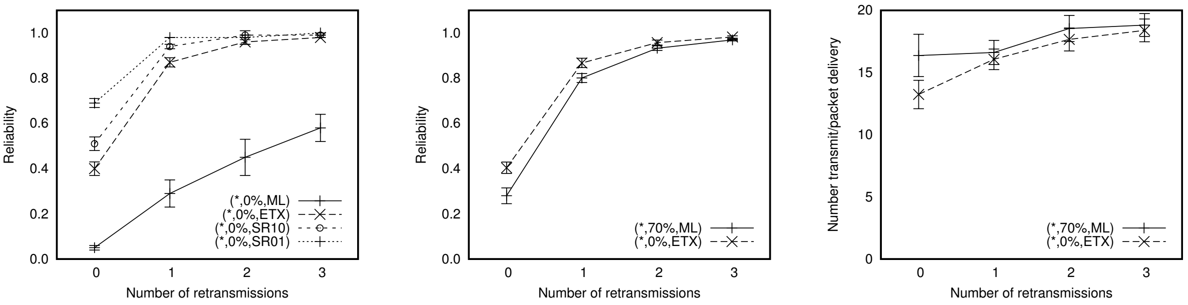
Each source generates one data packet every two seconds. Our simulations do not aggregate data within the network, exercising our reliability mechanisms more than they might have been in a system with in-network aggregation. We set an interest epoch of 100 seconds, so any routing choice affects about 50 data packets.

Our simulation results report 95% confidence intervals for each metric, obtained from 20 simulation runs.

B. Simulation Results

We have extensively explored the parameter space of interaction between blacklisting, reliability metrics and link-layer retransmission. For reasons of space, however, this section is organized in a manner that brings out the main results. Recall that the focus of our explorations is to find a set of mechanisms that enables highly

[1] A threshold of 0 does not filter out any bad links and the behavior of the underlying routing protocol is unchanged.

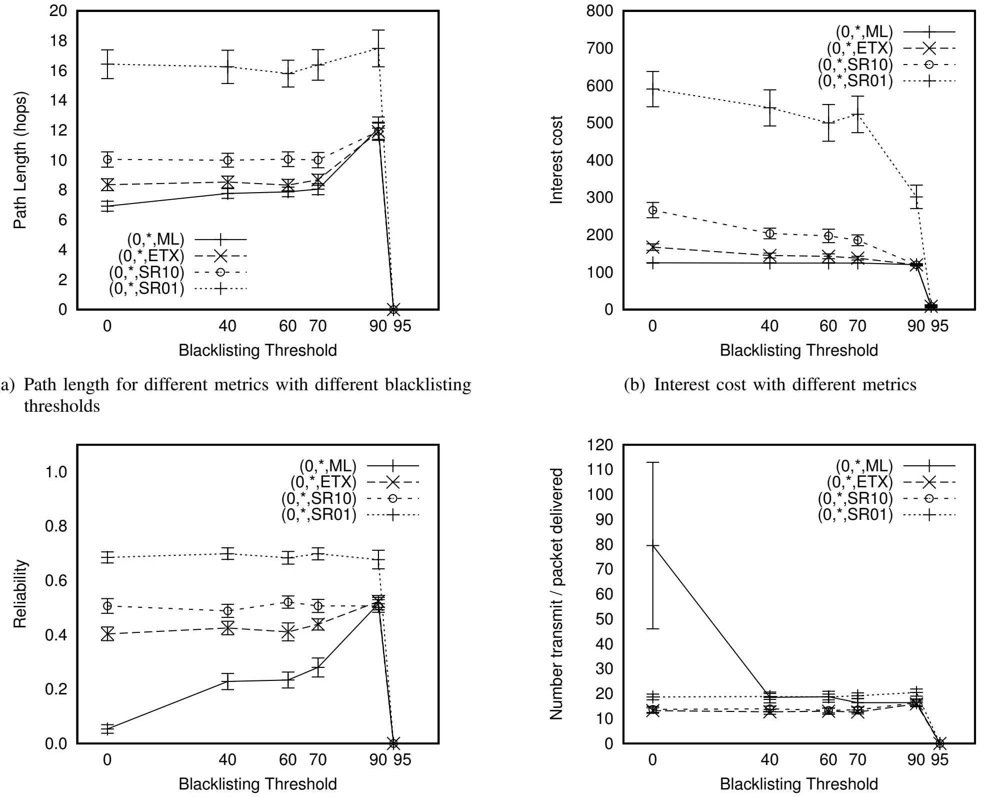


(a) Reliability with 0-3 retransmissions

(b) Reliability for (*,70%,ML) and (*,0%,ETX)

(c) Delivery Cost for (*,70%,ML) and (*,0%,ETX)

Fig. 2: Evaluation of retransmission alone, and in combination with blacklisting and reliability metrics



(a) Path length for different metrics with different blacklisting thresholds

(b) Interest cost with different metrics

(c) Reliability with different blacklisting thresholds

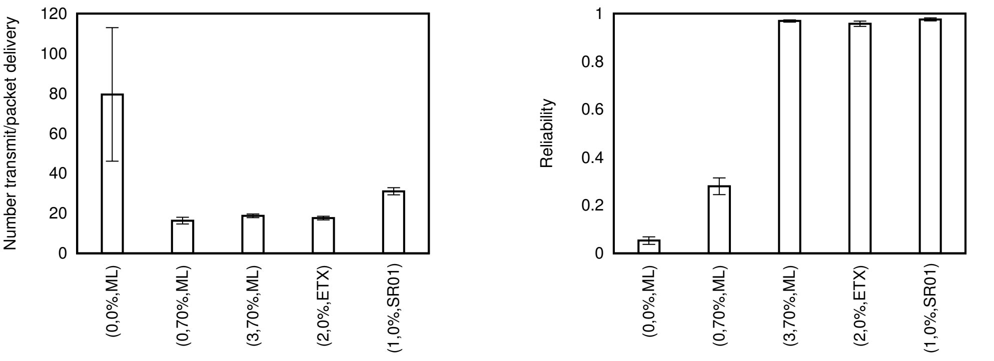
(d) Delivery cost at different blacklisting thresholds

Fig. 3: The interaction of blacklisting and routing metrics without any retransmission

reliable delivery with low overhead. We base our initial discussions on simulation of a network with an average density of 23 neighbors having a non-zero reception rate. Towards the end of the section, we discuss the impact of density.

Our first result is that link-layer retransmissions are *necessary* for achieving high reliability, given the packet loss rates observed in practical sensor network settings. Figure 2(a) shows that, without retransmissions, none of our mechanisms exhibit path reliability exceeding 70%. In addition, Figures 2(a) and 2(b) show that a small number of retransmissions is *sufficient* to achieve high path reliability (above 90%) when used in combination with either blacklisting or a reliability metric. That retransmissions can improve path delivery is somewhat obvious, but it is worth emphasizing particularly because commonly used sensor network MAC (such as those in TinyOS) either omit ARQ or make it optional.

High reliability can come at the cost of high overhead, however, if path lengths become too long or number of



(a) Data delivery cost with retransmission, blacklisting, and ETX

(b) Reliabilities with retransmission, blacklisting, and ETX

Fig. 4: Comparison of leading protocol combinations

retransmissions excessive. Next, we observe that ETX together with retransmissions can achieve high reliability *efficiently*. Figures 2(b) and 2(c) show that ETX can achieve nearly 98% delivery with up to 3 retransmissions with about 18 transmissions per delivered packet. Since typical path lengths incurred for ETX are 8–9 hops, this suggests about two transmissions per hop.

The efficiency of metric-based routing can depend on the choice of metric. Figure 2(a) shows that a higher resolution reliability metric (SR01) achieves a higher reliability (68%) than lower resolution reliability metrics (ETX at 40% and SR10 at 51%) at zero retransmissions. It would then seem that tuning the resolution of the reliability metric can improve reliability significantly. However, doing so increases path length and overhead. With a high-resolution metric (SR01), paths are twice as long as the ML metric (Figure 3(a) at 0% blacklisting threshold); high-resolution metrics are clearly unacceptable for latency critical applications. Similarly, a high resolution metric triggers many more route updates; Figure 3(b) shows that, without blacklisting, SR01 incurs 591 transmissions (4.7 transmissions per node) during interest propagation, more than twice that of ETX or SR10. For this reason, ETX together with a small number of retransmissions provides better path selection at low overhead. We note that this result, while not startling, is new: ETX has been shown to have good performance, but, to our knowledge, its performance in concert with link-layer retransmissions had not been studied before.

More surprising is the observation that the ML metric, together with blacklisting and retransmissions is able to achieve comparable reliability *at lower overhead* than ETX with retransmissions (Figure 4(a)). Essentially, blacklisting enables ML to find paths that have moderate reliability, and retransmissions on these paths improves path reliability (for example, compare the differences in reliability between (0,0%,ML), (0,70%,ML), and (3,70%,ML) in Figure 4(b)). ETX pays a higher interest overhead to find similar, high-reliability paths. Thus, Figure 3(b) shows that interest overhead for ETX is 33% higher than that for the ML metric, and arises from the fact that metric-based routing must propagate route updates as higher-reliability paths supersede low-latency, low-reliability paths. Blacklisting, on the other hand, immediately rejects these paths as beneath threshold.

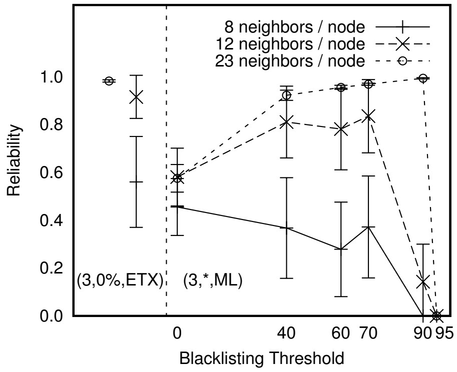


Fig. 5: The interaction between density and blacklisting. (3,0%,ETX) shown for comparison.

Variations in network density can affect these results. Unfortunately, while (3,70%,ML) is comparable to (3,0%,ETX) at high densities, the reliability of blacklisting falls off at lower density deployments. For example, as Figure 5 shows, at 12-neighbor density, (3,70%,ML) is about as reliable as (3,0%,ETX) (83% vs. 92%) with 30% lower interest costs. But when we consider less dense deployments in Figure 5, the reliability of (3,70%,ML) falls off because blacklisting begins to partition the network, rejecting unreliable but necessary links. For these reasons, we conclude that despite its higher interest cost, ETX together with retransmissions

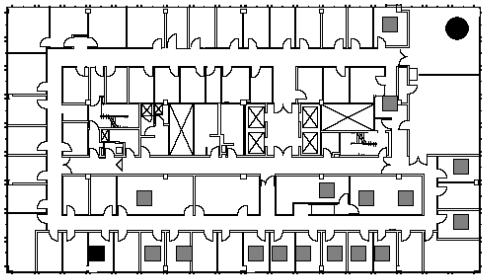


Fig. 6: Testbed deployment map. Gray boxes are relay nodes. The black circle (top-right) is a source node. The black box (bottom-left) is a sink node.

is the most desirable alternative since it is more stable across a range of densities. However, in a high density deployment of a sensor network, blacklisting may be preferable because of its lower interest cost.

The disadvantage of ETX is its higher interest cost. We hypothesized that the addition of moderate blacklisting to metric-based routing could serve to reduce this cost. Simulation results do not support this hypothesis for ETX, but (as we show later) our testbed results do. In simulation, Figure 3(b) shows, interest overhead for (3,*,ETX) is fairly constant with moderate blacklisting values (0–90%). A similar observation is true for path reliability (Figure 3(c)) and delivery cost (Figure 3(d)). Testbed results reach a quite different conclusion, suggesting this topic as an area for future work.

VI. Testbed Experiments

To validate our simulation results we conducted experiments on an 18-node testbed. Given the logistical difficulty of exploring the entire space of 96 experimental configurations, we chose five configurations as representative samples. Of these, the configuration (0,0%,ML) forms the baseline, (1,60%,ETX) has all three mechanisms (retransmissions, blacklisting, and a reliability metric), and (1,60%,ML), (0,60%,ETX), and (1,0%,ETX) consider combinations of two mechanisms each at one specific parameter setting.

A. *Methodology*

In our 18-node Stargate [10] testbed, we configured one node to function as a source, one as a sink, and 16 nodes as relays. Figure 6 is a map that shows how these nodes are deployed on a floor of our office building. A Mica-2 node attached to each Stargate was used for radio communication. We adjusted the radio transmit power on the mote such that each node has 5-15 neighbors. This setting provides a rich network connectivity (Figure 7) which makes available numerous possible paths between the source and the sink. The motes run TinyOS, but with S-MAC [20] as the MAC layer.

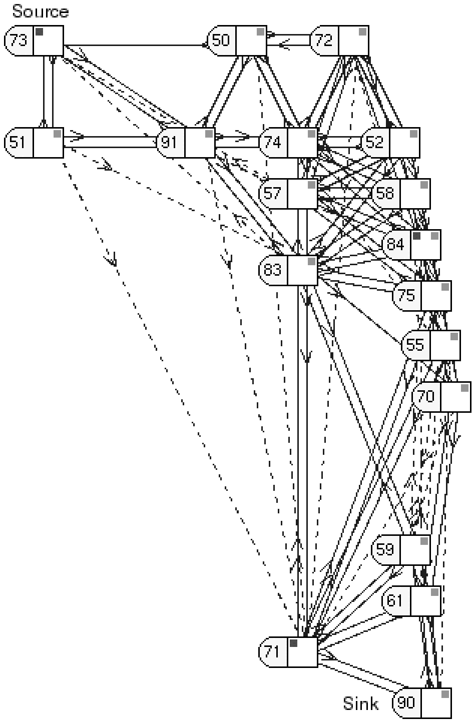


Fig. 7: Network connectivity in the testbed. Dotted lines indicate links with less than 60% reliability in one direction.

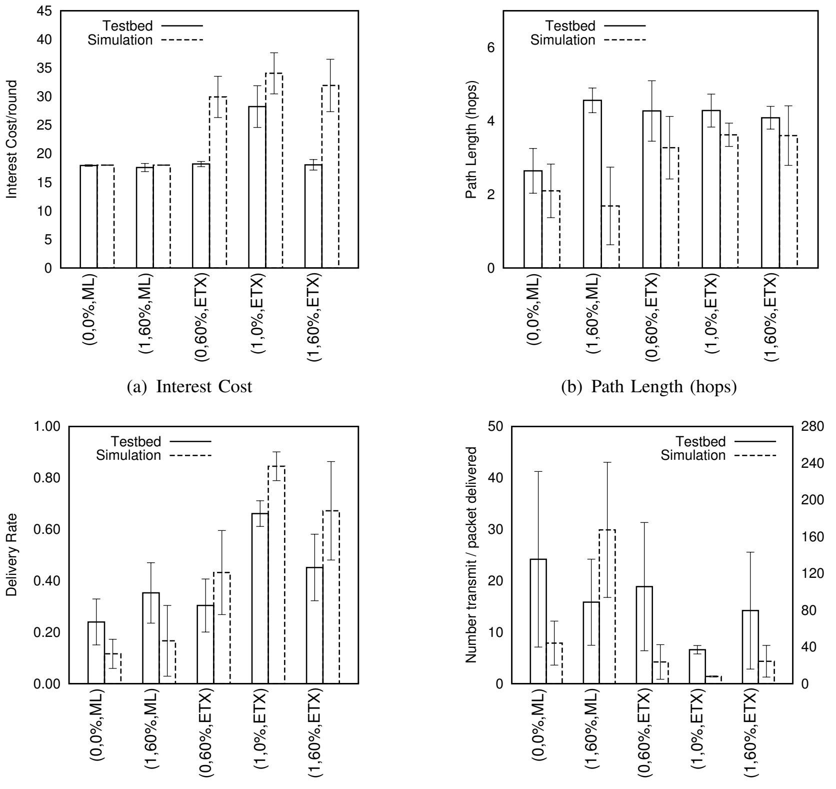
We used Emstar [8] on each node to setup and control the experiments. Emstar on each Stargate node uses the Mica-2 as a network interface. We use Emstar's link statistics collection module and its blacklisting module. Emstar's link statistics collection uses WMEWMA to estimate link quality to its neighbors. Its blacklisting module uses the link statistics estimate to identify links that have delivery rate below a configured threshold and disables those links.

Minimal software modification was necessary to the simulation software to run it on the Stargate testbed. Diffusion was configured to use Emstar's blacklisting capability, and to obtain link statistics from Emstar. We configured Emstar to send neighbor probes every 10 seconds. In our experiments, each configuration ran for 37.5 minutes. During that time, the sink sent 15 rounds of interest and the source sent data every three seconds.

Finally, in order to validate our simulation results on the testbed, we collected the temporal link statistics and topology information from the testbed and *input those into the simulator*. The next section compares the results obtained from simulation, with results from our testbed.

B. *Results from Testbed Evaluation*

Figures 8(a) through 8(d) compare the values of different metrics obtained using the testbed and from a



(a) Interest Cost

(b) Path Length (hops)

(c) Delivery Rate

(d) Delivery Cost

Fig. 8: Comparison of Testbed and simulation results for various configurations using four metrics

comparable simulation, for the five configurations. There is, by and large, remarkable agreement between our testbed experiments and our simulation: for most metrics and for most configurations, the difference between experiment and simulation falls within the bounds of experimental error. This gives us confidence that our conclusions (Section V) will hold in practice. For brevity, we focus on situations where there is some disagreement between experiment and simulation.

In many cases, the testbed results are slightly different from those obtained using simulations on the *same* topology. Lacking a detailed instrumentation at the MAC layer, we are not able to isolate the cause for this discrepancy. We conjecture this difference can be explained by observing that the simulator may not accurately capture interference from concurrent transmissions. Furthermore, the testbed results exhibit greater variability than simulations on the same topology. This can be attributed to fewer number of nodes and runs on the testbed. We plan to verify our experimental results on a larger testbed.

Results from the testbed also have a higher variability relative to those discussed in Section V, particularly for configurations with blacklisting, We attribute this to our earlier simulations' not capturing the temporal variations in link quality observed in the testbed.

Finally, one configuration where experiment deviates from simulation is the impact of blacklisting on interest overhead when used in conjunction with ETX. Figure 8(a) shows that (1,0%,ETX) uses about 28 total messages in a 18-node network while the ML metric only uses about 18 total messages every interest epoch. This contradicts our simulation results, which suggest that blacklisting has a negligible effect on reducing interest overhead. We do not have an explanation for this discrepancy at the time of writing.

VII. Conclusion

In this paper, we have examined the interplay between three mechanisms for selecting highly reliable wireless routing paths at low overhead: blacklisting, reliability metrics, and retransmission. To our knowledge, ours is the first work to systematically evaluate this design space. Our simulations reveal several interesting results: link-layer retransmissions are necessary for high path reliability; a reliability metric like ETX, together with up to three link-layer retransmissions can provide high path reliability at low overhead; more surprisingly, the ML metric together with blacklisting and retransmissions can often provide comparable reliability with slightly lower overhead, but this configuration is sensitive to the blacklisting threshold. Given these results, we recommend that a reliability metric such as ETX, together with link-layer retransmissions, is a robust choice that works well across the range of configurations we explored. The remarkable agreement between simulation and a real-world testbed lends significant weight to our conclusions.

REFERENCES

[1] Herbert Rubens Baruch Awerbuch, David Holmer. High throughput route selection in multi-rate ad hoc wireless networks. In *Proceedings of the First Working Conference on Wireless Ondemand Network Systems (WONS 2004)*, August 2004.

[2] Vaduvur Bharghavan, Alan Demers, Scott Shenker, and Lixia Zhang. MACAW: A media access protocol for wireless LAN's. In *Proceedings of the ACM SIGCOMM Conference*, pages 212–225, London, UK, September 1994. ACM.

[3] Kwan-Wu Chin, John Judge, Aidan Williams, and Roger Kermode. Implementation experience with manet routing protocols. *SIGCOMM Comput. Commun. Rev.*, 32(5):49–59, 2002.

[4] Douglas S. J. De Couto, Daniel Aguayo, John Bicket, and Robert Morris. A high-throughput path metric for multi-hop wireless routing. In *Proceedings of the 9th ACM International Conference on Mobile Computing and Networking (MobiCom '03)*, San Diego, California, September 2003.

[5] Douglas S. J. De Couto, Daniel Aguayo, Benjamin A. Chambers, and Robert Morris. Performance of multihop wireless networks: Shortest path is not enough. In *Proceedings of the First Workshop on Hot Topics in Networks (HotNets-I)*, Princeton, New Jersey, October 2002. ACM.

[6] Richard Draves, Jitendra Padhye, and Brian Zill. Comparison of routing metrics for static multi-hop wireless networks. In *Proceedings of the ACM SIGCOMM Conference*, Portland, OR, USA, August 2004. ACM.

[7] R. Dube, C. Rais, K. Wang, and S. Tripathi. Signal stability based adaptive routing (ssa) for ad hoc mobile networks, 1997.

[8] J. Elson, S. Bien, N. Busek, V. Bychkovskiy, A. Cerpa, D. Ganesan, L. Girod, B. Greenstein, T. Schoellhammer, T. Stathopoulos, and D. Estrin. Emstar: An environment for developing wireless embedded systems software, 2003.

[9] John Heidemann, Fabio Silva, and Deborah Estrin. Matching data dissemination algorithms to application requirements. Technical Report ISI-TR-571, USC/Information Sciences Institute, April 2003.

[10] Crossbow Technology Inc. Stargate platform. http://www.xbow.com/Products/XScale.htm.

[11] Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, John Heidemann, and Fabio Silva. Directed diffusion for wireless sensor networking. *ACM/IEEE Transactions on Networking*, 11(1):2–16, February 2002.

[12] Hang Liu, Hairuo Ma, Magda El Zarki, and Sanjay Gupta. Error control schemes for networks: An overview. *Mobile Networks and Applications*, 2(2):167–182, 1997.

[13] Henrik Lundgren, Erik Nordstro, and Christian Tschudin. Coping with communication gray zones in ieee 802.11b based ad hoc networks. In *Proceedings of the 5th ACM international workshop on Wireless mobile multimedia*, pages 49–55. ACM Press, 2002.

[14] Yogesh Sankarasubramaniam, Özgür B. Akan, and Ian F. Akyildiz. ESRT: Event-to-sink reliable transport in wireless sensor networks. In *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing*, page xxx, Annapolis, Maryland, USA, June 2003. ACM.

[15] Fred Stann and John Heidemann. RMST: Reliable data transport in sensor networks. In *Proceedings of the First International Workshop on Sensor Net Protocols and Applications*, pages 102–112, Anchorage, Alaska, USA, April 2003. IEEE.

[16] Chieh-Yih Wan, Andrew Campbell, and Lakshman Krishnamurthy. PSFQ: A reliable transport protocol for wireless sensor networks. In *Proceedings of the ACM Workshop on Sensor Networks and Applications*, pages 1–11, Atlanta, Georgia, USA, September 2002. ACM.

[17] Chieh-Yih Wan, Shane B. Eisenman, and Andrew T. Campbell. Coda: Congestion detection and avoidance in sensor networks. In *Proceedings of the First ACM SenSys Conference*, pages 266–279, Los Angeles, California, USA, November 2003. ACM.

[18] Alec Woo, Terence Tong, and David Culler. Taming the underlying challenges of reliable multihop routing in sensor networks. In *Proceedings of the first international conference on Embedded networked sensor systems*, pages 14–27. ACM Press, 2003.

[19] Mark D. Yarvis, W. Steven Conner, Lakshman Krishnamurthy, Alan Mainwaring, Jasmeet Chhabra, and Brent Elliott. Real-world experiences with an interactive ad hoc sensor network. In *Proceedings of the Internal Workshop on Ad Hoc Networking*, August 2002.

[20] Wei Ye and John Heidemann. Medium access control in wireless sensor networks. Technical Report ISI-TR-580, USC/Information Sciences Institute, October 2003. To appear as a chapter in *Wireless Sensor Networks*, Taieb Znati, Krishna M. Sivalingam and Cauligi Raghavendra (eds.), Kluwer Academic Publishers.

[21] Jerry Zhao and Ramesh Govindan. Understanding packet delivery performance in dense wireless sensor networks. In *Proceedings of the first international conference on Embedded networked sensor systems*, pages 1–13. ACM Press, 2003.