Do Sensornet Protocol Variants Yield Real Benefits?

Jonathan Tate Iain Bate
Department of Computer Science, University of York,
York, YO10 5DD, United Kingdom
{*jt* | *iain.bate*}*@cs.york.ac.uk*

Abstract

Simple, stateless networking protocols provide a low-cost and predictable foundation upon which to build decentralised applications. Sensornets are complex, containing emergent behaviour; identifying protocols offering appropriate behaviour may be difficult. In this paper we evaluate the relative performance of gossiping protocol variants in non-ideal sensornets. We examine the extent to which a "standard" gossiping protocol might be outperformed by variants of this standard, each specialised and optimised to mitigate anticipated problems. We categorise and measure the undesired behaviours that remain, as a foundation for future protocols which will address these specific issues.

1 Introduction

Sensornets, an emerging discipline of embedded system and network design, are distributed network systems composed of many low-cost independent nodes. Each node is embedded into a physical environment of interest, and is equipped with sensors with which to extract raw data about this environment. Distributed sensing applications require data from numerous physical locations to synthesise conclusions about the environment.

A consistent thread running through sensornet research is that they must function effectively under tight resource constraints. Nodes have restricted energy, computation and storage resources and therefore limited utility in isolation; cooperation and coordination is necessary to address realistic problems. Whether implemented using specialised discrete motes, or as functionality piggy-backed on existing equipment, it is vital that resource usage is kept to an absolute minimum. This maximises the lifetime of networks whose power sources cannot be replenished [15], and minimises the cost of necessary hardware.

It is equally important that the sensornet keeps pace with the real world with which it interacts, and has sufficient redundancy to recover from individual node failures. Contradictory requirements lead us to realise that multi-objective optimisation is essential to ensure that a reasonable compromise can be found [25]. Furthermore, some protocol behaviours become evident only when assessed in systems of realistic scale [26], which may be larger than available physical testbed networks. Studies examining protocol behaviour at sufficient scale to provide confidence of verity and applicability remain scarce to date [18].

A common strategy is to design new protocols for a specific intended deployment context [21]. We believe this strategy is flawed. Real-world deployment environments are usually not known precisely prior to deployment, and often change after deployment. Without detailed knowledge of specific difficulties faced in a given environment, and specific weaknesses observed when attempting to apply a given protocol, there is little evidence that custom protocols effectively address any real requirement. Custom protocols often lack generality and become unusable if the expected and actual deployment environments differ, and may have unexpected behaviours or failure modes which would have been previously observed in commonly used protocols.

An alternative strategy is to first tune existing protocols to application requirements, only moving on to custom protocols if no possible configuration yields acceptable results. Tuning gossiping protocols to specific deployment scenarios is explored in [25]. The multi-objective optimisation problem is addressed using an engineering method based on the Design of Experiments [2] approach. Although this yields useful and directly deployable results, it offers little significant insight into the underlying problems.

In this paper we also consider network behaviour within a given sensornet deployment. We do not attempt to define new protocols, or to configure existing protocols, so as to perform optimally in this scenario. Instead, we define and classify specific types of undesired or suboptimal behaviours which may adversely affect sensornet performance, and measure the occurrence of these under a set of gossiping protocol variants. This provides both a qualitative and a quantitative foundation for the design or customisation of protocols to address these real, specific weaknesses.

The structure of the remainder of this paper is as follows. Section 2 discusses related work. Section 3 describes the use of simulation experiments. Section 4 measures systems implementing a "standard" gossiping protocol, proposing and evaluating protocol modifications to address observed weaknesses. Finally, section 5 presents a summary.

2 Related work

Flooding is perhaps the simplest possible *network routing protocol*. Upon generating a traffic packet a node broadcasts this to each of its neighbours. Each of these neighbours then rebroadcasts the packet to each of its neighbours, and so on, the process converging in $O(d)$ rounds in network diameter d where almost all nodes have been exposed to the packet [16], with some exceptions [10]. Most non-geographical routing protocols use flooding in some capacity, often for initial route discovery [12].

Counterintuitively complex behaviour is observed [10] despite the protocol's simplicity. *Broadcast storms* [19] are particularly problematic with significant redundant broadcasts, contention, collisions, and high energy consumption. Poor scalability in network size is a major weakness. Nevertheless, it delivers packets with high probability and low latency. Being stateless in nature it requires few resources and requires no network knowledge, and as such often succeeds in rapidly changing networks where more sophisticated protocols struggle [20].

Gossiping extends flooding by having all nodes rebroadcast with probability $p \in [0,1]$. Higher values of p give higher probabilistic delivery guarantees. Bimodal behaviour in p is observed where packets reach either very few or very many nodes, with sharp transition about some critical probability p_c where often $p_c \in [0.6, 0.8]$ [12]. Selecting p is generally difficult. Typical goals include improving delivery probability by preventing *premature gossip death* [12]. A myriad of variants of flooding and gossiping protocols have been proposed but to the best of our knowledge none are based on measurements of the defective behaviours, through detailed modelling or experimentation, for which they are intended to compensate.

Ni *et al.* [19] propose probabilistic, counter-bounded, distance-based and location-based schemes, but do not combine probability with other factors. *GOSSIP1* [12] sets $p = 1$ for the first few hops before reverting to the standard network-wide p. *GOSSIP2* [12] extends *GOSSIP1* by raising p to $p_a > p$ at nodes with few neighbours, which is particularly beneficial in sparse networks.

Barrett *et al.* [4] define the *destination attractor* concept in which p is decided per broadcast and increases as the packet approaches the destination, and *directed transmission* in which p increases in proportion to the number of hops already travelled. These ideas are implemented in *PPSNRP*, but measure distance-to-destination in network hops which is often unknown and in any case may be an inappropriate measure in unevenly distributed networks.

GeRaF [28] applies a similar principle to the *destination attractor* of *PPSNRP* but uses physical node locations rather than hopcounts. This removes the unrealistic requirement of maintaining hopcount-to-destination data at all nodes, but requires sending nodes to know the physical location of the intended destination(s). Li *et al.* [17] instead use geographic information to define bounding ellipses beyond which rebroadcast activity is forbidden.

PGR [23] encourages delivery along the straight line defined by S and D. When B broadcasts it assigns rebroadcast probability p_i for each neighbour N_i as a function of the angle θ_i between the vectors $\overrightarrow{BN_i}$ and $\overrightarrow{BD}$. Smaller angles θ_i tend to produce shorter and more efficient routes, and are assigned higher p_i. *IGF* [13] implements a similar strategy but allows neighbours N_i to decide independently whether to rebroadcast in a three-phase protocol.

Optimising sensornets for energy efficiency is complex. Raghunathan [22] observes that it *involves not only reducing the energy consumption of a single sensor node but also maximising the lifetime of an entire network*, requiring dynamic tradeoffs between *energy consumption*, *system performance*, and *operational fidelity*, yielding *up to a few orders of magnitude* of improved lifetime. Protocol selection plays an important role in any comprehensive solution.

To achieve an appropriate QoS it is necessary to consider each non-functional requirement throughout the protocol stack [22], influencing the design and operation of sensornet applications, networking protocols, network topologies and network tasking. Grenier and Navet [11] illustrate MAC protocol fine tuning to achieve certain real-time requirements while conforming to other application-dependent criteria, but non-*Data Link Layer* issues are ignored.

With many controlled factors and measured responses it is generally difficult to understand the resulting complex interrelationships. Totaro and Perkins [27] apply a *systematic statistical Design Of Experiments* approach to evaluate and model the complex tradeoffs. This work considers the impact of controlling network composition with a fixed network application and environment. However, the underlying protocol weaknesses are not explored by this method.

Simulation offers an environment in which experiments are perfectly repeatable and perfectly controllable. This is generally impossible in real world experiments, and is particularly acute in sensornet research where experiments must consider interaction with non-network entities. Anastasi *et al.* [1] state that *simulation is clearly a better choice than experiments when considering large networks, as controlling large testbeds is very hard*, and also state that *most of the research on ad hoc networks has been carried out by adopting simulation and analytical approaches only*.

3 Measurement by simulation experiment

Sensornets are an emerging technology which has already enjoyed some commercial success. However, deploying a sensornet in real-world applications remains a difficult, expensive, and error-prone challenge [3]. A consequence of this difficulty, and the cost of mote hardware [6], is that few sensornets of very large scale have been deployed into real environments. Despite the abundance of high quality protocols described and examined in the literature, no single sensornet routing protocol has yet attained the status of a *de facto* standard, as one might consider IP a *de facto* standard in commodity wired networking [8].

As sensornets are inherently application-specific [14] it is possible that a similar level of homogeneity is never attained. This is not necessarily a negative observation. If a sensornet is to be deployed into a specific environment, and is to be a self-contained system which does not interoperate with other sensornets, the network designer is free to select the most appropriate hardware, middleware and application software components and optimise the composition without regard for generality. Nevertheless, it is generally desirable to reuse existing network components rather than create new examples, unless no suitable candidates exist.

We therefore select a class of protocol known as the *gossiping* protocols [19, 12]. These are designed for sensornet-like deployment contexts, and have been examined thoroughly in the literature. These protocols are *stateless* and therefore do not consume storage resources to maintain routing tables or details of the underlying network. Furthermore, these protocols do not incur the energy and bandwidth overheads associated with distributing and maintaining this information, and do not suffer from problems arising from expired or redundant information. These issues are particularly relevant in highly dynamic or unreliable sensornets. These protocols have low computational complexity and require little working space in memory, and are therefore ideal for sensornets composed of motes with highly restricted computational and energy resources.

The methods described in this paper can be reapplied to any arbitrary class of protocol variants, but it is logical to first consider the simple case. In selecting the class of *gossiping* protocols we make no claims as to their merit for any given sensornet application. More specifically, we do not claim that when optimally configured they necessarily offer superior performance to other recent and more complex alternative protocols. However, their simplicity implies they are readily understood and analysed, and it is easy to make modifications to generate new variants which are consistent with existing known behavioural properties.

Some interesting effects and behaviours may only become evident in sufficiently large networks. Preliminary simulation experiments considered networks of variable numbers of similar nodes distributed with constant spatial density [26]. Qualitatively different behaviour was observed in networks of 200 nodes and of 500 nodes, with additional features and points of inflection appearing in plotted curves for the latter. Increasing node count further to 750, 1000 or 2000 nodes did not yield further features. We conclude that a test network size of 500 nodes is sufficient.

Measurement of network behaviour influenced by protocol selection, configuration, and customisation, would ideally occur in physical testbed networks of realistic scale and composition. Unfortunately, economic and logistical factors generally preclude the construction of test networks on the order of hundreds of nodes for such experiments. Physical testbed networks of this scale would be unable to provide sufficient repeatability to extract meaningful results. Simulation provides a solution to these problems. Large sensornet systems can be evaluated, at acceptable cost, in repeatable and deterministic environmental conditions.

The accuracy of simulated environments is a significant concern when obtaining experimental results. All experiments in this paper were implemented using the *yass* [24] sensornet simulation tool. Previous work with a gossiping protocol and *yass* demonstrates that significant variation in sensornet performance is observed for different protocol configurations. It follows that similar variation in measured sensornet performance should be observed if some variants of gossiping protocols yield significant benefit.

Three simulated networks were defined, each of 500 nodes. Each network was identical in all regards other than node spatial distribution. By averaging or otherwise compositing results across these three test networks we ensure that the characteristic quirks of any given network do not exert undue influence. Node spatial distribution within a bounding volume is random and even, with constant spatial density of 1.5×10^{-7} *node* m^{-3}. Simulated nodes were based on the MICA2 sensornet mote [7] with radio range of around 150m, although this detail is largely irrelevant as any similarly-equipped nodes will yield similar behaviour.

In the simulated application each node serves as a packet source and packet sink, utilising the unicast paradigm throughout. Each node generates packets sporadically with a single randomly selected destination to model a general distributed and decentralised process control application. Simulated packets have length randomly selected in the interval $[128, 1024]$ bits, including the header. With the MICA2 radio having a transmit speed of 38.4*Kbps* [7] this gives per-packet transmit times in the interval $[3.33 \times 10^{-3}, 2.67 \times 10^{-2}]$ seconds.

When packet transmission begins the local wireless medium is occupied for some duration in this interval. Nodes implement CSMA such that, if attempting to broadcast a packet, an exponential backoff procedure is implemented should the nearby wireless medium be occupied. A

waiting node will implement up to 8 sense-wait cycles, doubling the wait period on each iteration, before giving up and dropping the packet. Note that although this greatly reduces packet broadcast collisions it does not avoid the *hidden terminal* problem [9] inherent in wireless communications.

4 Online dynamic protocol tuning

In this section we consider the performance of gossiping protocol variants which dynamically modulate their behaviour in response to observed network conditions, packet delivery attempt status, or other measurable factors, in an attempt to improve network performance. The fundamental concept is that individual packets may receive significantly different treatment compared to that received by their peers, within the framework of a protocol which applies rules consistently for all packets.

We define a set of metrics in section 4.1, to identify failure modes of the protocol. By examining the frequency with which values of these metrics are observed we quantitatively characterise the conditions under which packet delivery attempts fail. These metrics were chosen because all but one can be determined online within the network at minimal cost, and hence can be used to dynamically tune protocols to improve routing decisions with minimal overhead.

Each simulation was allowed to continue for 3600 simulated seconds to allow network behaviour to settle into stable patterns. A single gossip protocol configuration was taken as a baseline with static gossip rebroadcast probability of p throughout the network. We set $p = 0.6$ as [26] demonstrates this is sufficient to ensure delivery of most packets to most nodes at most times in the experimental configuration defined in section 3, following the bimodal behaviour effect predicted by *percolation theory* [12].

The frequency distribution for each failure characteristic metric was determined by taking the output from three simulations, corresponding to the three networks described in section 3, and counting the frequency of values falling into equal-width bins spaced evenly throughout the observed range. Experimentation showed that 20 bins provided a good balance between detail and clarity in the derived histograms, exposing trends without too much distracting noise [5].

Graphs were plotted showing bin midpoint values versus relative frequency. Each plot summarises the observed behaviour in a given response type over the total simulated period, and approximates the probability distribution function of the response value for any given individual packet similar to the greater packet population [5].

For each measured attributed described in section 4.1 a pair of graphs illustrates the approximated distribution function. The first graph shows the distributions for all packets, delivered packets, and undelivered packets. This graph illustrates undesirable behaviours evident for undelivered packets which differ from those of successfully delivered packets. The second graph shows the distribution of undelivered packets under variants of the gossip protocol, modified to address the suboptimal behaviour displayed in the first graph.

Each attribute is measured with respect to the *closest delivery attempt node*, C. We assume each packet π has a single source node, S, and a single destination node, D. For delivered packets, trivially $C = D$. For undelivered packets, C is the node geometrically closest to D which successfully received a copy of π.

Figure 1 illustrates the distribution of distances from *source* to *closest*, $||\overrightarrow{SC}||$. Figure 2 illustrates the distribution of distances from *closest* to *destination*, $||\overrightarrow{CD}||$. Figure 3 illustrates the distribution of hopcounts encountered by the first instance of packet π reaching C. Figure 4 illustrates the distribution of time elapsed during which the first instance of packet π reaches C.

Figure 5 illustrates the distribution of total numbers of node-to-node hops observed in delivery of a packet. This latter figure is taken as a heuristic measure of the total network resources expended in attempts to deliver a packet, assuming that each node-to-node hop consumes similar levels of resources of interest such as energy.

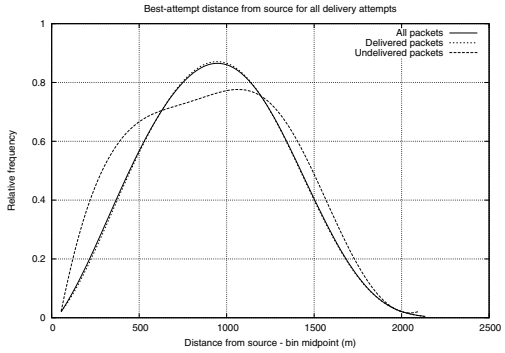
4.1 Distribution analysis: unmodified protocol

In this section we assess qualitatively the characteristic behaviour observed for packets generated by the simulated application. The behaviour demonstrated for all packets is compared against that demonstrated by those subsets which were delivered, and those not delivered. We assume that qualitative differences between the resulting distributions represent qualitative differences in behaviour of packets which do and do not reach their destination.

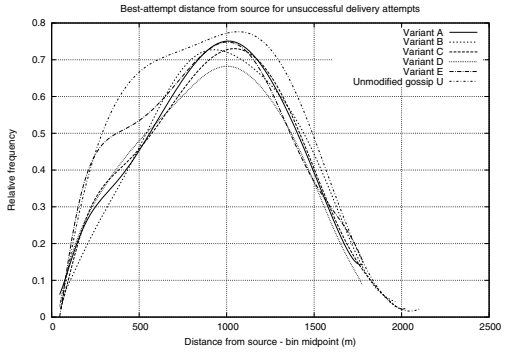
Accepting that some level of packet loss is inevitable in a wireless network not designed or optimised for reliability, ideally the loss probability would be equal for all packets as the simulated networks treat all packets as being of equal priority. Suppose some qualitative difference exists; this represents a weakness in the behaviour induced by a given protocol, which might be addressed by amending the protocol to deal with the specific cases in which this weakness becomes influential.

The key problematic behaviours associated with failed packet delivery attempts are *early fail* and *late fail*, corresponding to the bimodal behaviour inherent to gossiping protocols [12]. In *early fail* gossiping dies out early, wasting few resources, and stifling the delivery attempt before it becomes established. In *late fail* the packet covers most of the network and yet fails to reach the destination, thereby wasting all resources invested in the attempt.

In figures 1(a)-5(a) it is evident that for all metrics the probability distribution function for all packets is almost identical to that for delivered packets. This is unsurprising as around 92% of packets were successfully delivered in each case. Refer to [26] for full experimental details. Furthermore, the distribution function for undelivered packets is generally similar to that of all or delivered packets, albeit with some difference in overall shape. It is these differences in shape we shall examine here. The relative height of features in the plots is not significant, as all have been scaled in the *y*-axis such that discrimination of individual plots is possible.



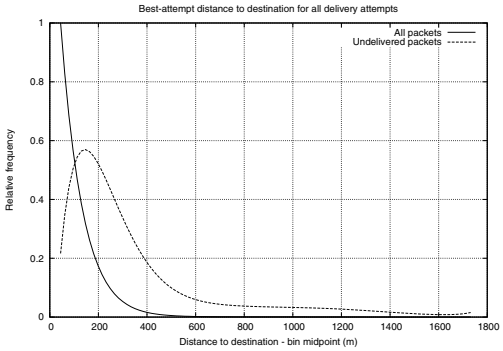
(a) All delivery attempts for unmodified gossip



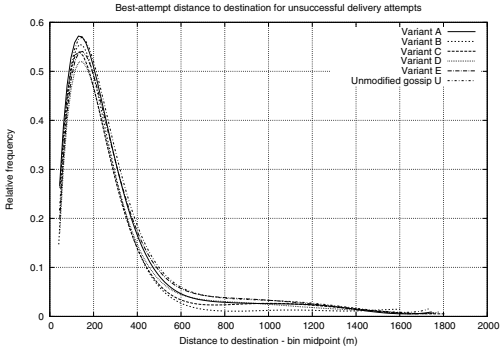
(b) Unsuccessful delivery attempts for gossip variants

Figure 1. Source distance vs frequency

Figure 1(a) shows the distribution of relative frequency of distances to observe a good approximation of a Gaussian distribution. This is an expected geometric consequence of randomly selecting source and destination nodes for traffic packets which are distributed randomly and evenly throughout a cube. The distribution pertaining to undelivered packets more roughly approximates a Gaussian distribution with similar mean, but with the distinct asymmetry that the falloff is much less pronounced below this mean than above. This shows that undelivered packets are substantially less likely to get close to their intended destination than delivered packets, providing evidence of the *early fail* of gossip message propagation [12].



(a) All delivery attempts for unmodified gossip



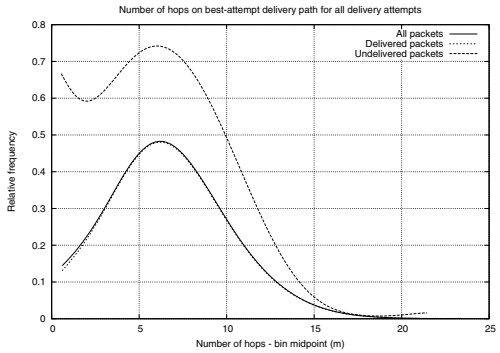
(b) Unsuccessful delivery attempts for gossip variants

Figure 2. Destination distance vs frequency

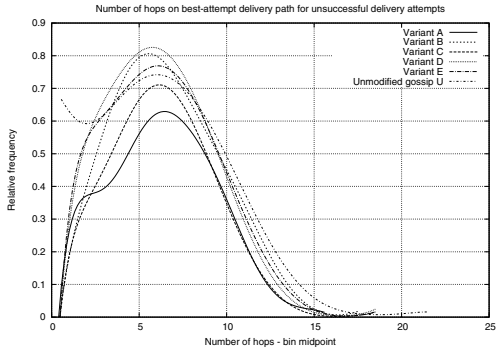
Figure 2(a) does not show the frequency distribution for delivered packets, as this is trivially zero at all places except at the origin where a Dirac delta would represent all delivered packets. In the distribution plot for all packets we find further evidence of bimodal behaviour in the hyperbolic curve; any given packet tends to cover almost all of the network as $p >= p_c$ in the experiment design and hence reaches nodes close to the intended destination.

The plot for undelivered packets tells much the same story, although there is a sharp falloff in frequency at the *y*-axis due to the discrete categorisation of packets as *delivered* or *undelivered*. Even undelivered packets tend to reach nodes close to the destination, representing wasteful employment of network resources; if a delivery attempt is to fail, it is better that failure occurs sooner rather than later, such that substantial energy and time need not be wasted when the packet falls just short of the target. This is the *late fail* effect we discuss above.

Figure 3(a) indicates that the path from the source node to whichever node of those receiving the packet is closest to the destination is far more likely to consist of a very small number of node-to-node hops for undelivered packets than for delivered packets. Of course, some source-destination

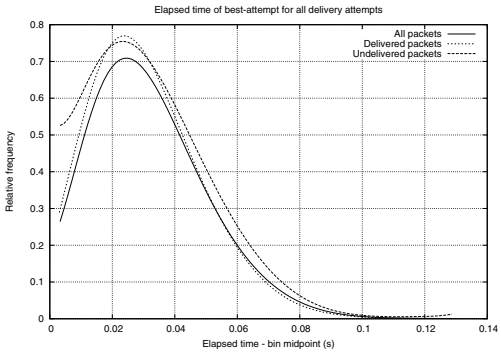


(a) All delivery attempts for unmodified gossip

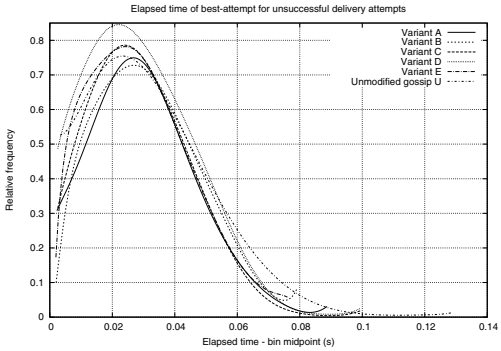


(b) Unsuccessful delivery attempts for gossip variants

Figure 3. Best-path hops vs frequency



(a) All delivery attempts for unmodified gossip



(b) Unsuccessful delivery attempts for gossip variants

Figure 4. Best-path time vs frequency

node pairs happen to be located in close physical proximity; it would be expected here that the best internode route would contain few hops. More generally, however, we see further evidence of *early fail*.

In figure 4(a) we see a similar effect to that illustrated in figure 3(a) when we consider time since delivery attempts began, although the qualitative difference between the distributions for delivered and undelivered packets is less pronounced. Nevertheless, these plots indicate that a time-based approach might offer a similar remedy to *early fail* to that offered by a hopcount-based approach.

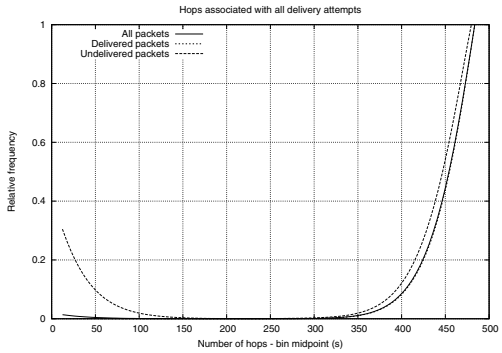
Figure 5(a) is slightly different to the others considered in this section. The plots represent a heuristic measure of energy consumed per packet for successful and unsuccessful delivery attempts, working from the assumption that each node-to-node hop consumes a roughly comparable quantity of energy available within the network. Again we see evidence of *early fail* behaviour. Under the unmodified gossip protocol we expect a packet to cover almost all, or almost none, of the network, regardless of the physical or logical proximity of the source and destination nodes.

If we assume that each node will rebroadcast a given packet only once then total network coverage in a network of n nodes would involve $n-1$ hops. A substantial area under the undelivered packets curve toward the left of figure 5(a) indicates a substantial number of packets for which very few hops are counted, indicating little of the network was exposed to the packet; unless by chance the destination happens to be physically close to the source this renders unlikely the successful delivery of the packet.

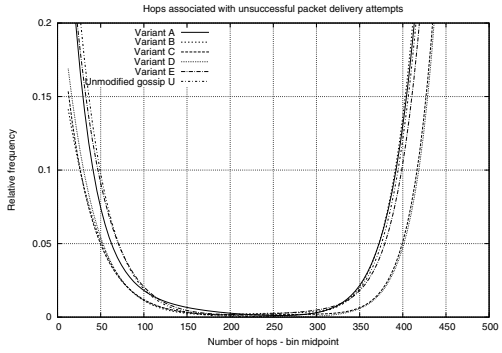
4.2 Distribution analysis: protocol variants

Having established and measured the undesirable characteristics of the unmodified gossip protocol associated with failed packet delivery attempts, we now attempt to attack these weaknesses. We modify the gossip protocol to yield the following variants in which online heuristics are applied, at each node and for each packet individually, to modify the effective rebroadcast probability p_e from the base probability p under certain circumstances.

The underlying concept is that we might improve network performance by encouraging rebroadcast in nodes or regions which are likely to form part of good delivery paths, and discourage rebroadcast where it is unlikely to make a net positive contribution. Nodes do not have complete knowledge of the network state or future traffic production, so an optimal algorithm is not possible; however, heuristic approaches might offer useful improvement approaching



(a) All delivery attempts for unmodified gossip



(b) Unsuccessful delivery attempts for gossip variants

Figure 5. Total hops vs frequency

theoretical bounds.

The gossip variants considered are:

- U: original unmodified gossip
- A: distance from source + elliptical bounding
- B: hops traversed + elliptical bounding
- C: time since start + elliptical bounding
- D: distance to destination + elliptical bounding
- E: elliptical bounding regions only

Variant U is the unmodified gossip protocol against which the variants are compared. Variant E implements the elliptical bounding of Li *et al.* [17], by setting $p_e = 0$ for nodes beyond a bounding sphere whose diameter is defined by the packet's source and destination nodes, to prevent packets "spilling out" into positions unlikely to contribute to successful delivery.

Variants $A-D$ also employ the elliptical bounding of E but augment this by increasing p_e under certain circumstances, provided that the packet is within the bounding region. Each gossiping protocol variant α of $A-D$ calculates p_e in the interval $[p,q]$ as a function of some measured packet attribute, s_α, where that attribute value falls between zero and some threshold value t_α. p represent the lower bound of p_e, and q represents the upper bound of p_e.

We define $q = 1$ to maximise the effect of the rebroadcast probability increase, although any $q \in [p,1]$ can be used. For $s_\alpha < t_\alpha$ the value of p_e is related to the value of s_α by the selected mapping function. For $s_\alpha = t_\alpha$ the value of $p_e = q$. For $s_\alpha > t_\alpha$ the value of $p_e = q$; no further increase in p_e is possible, as the defined upper bound has been reached.

The value of t_α for each variant $A-D$ was obtained from the graphs of section 4.1. This was implemented by identifying the range(s) of the metric, displayed on the x-axis, showing dissimilarity in the relative frequency of observations for delivered and undelivered packets. It is obvious that different threshold values may be appropriate in different sensornets. Variant E always has a unitless threshold value of 1, the relative length of source-to-intermediate and source-to-destination vectors, and hence is not measured.

Any number of functions might be selected, such as the *step function* or various types of *exponential* or *trigonometric* functions, but we select a simple *linear ramp function*. The latter requires few resources for calculation, and is a reasonable approximation of numerous functions of greater computational cost. For some variant α, the effective rebroadcast probability, p_e, for a given packet where the attribute measurement is s_α, is given by equation 1. This is a generalisation of this class of packet context-sensitive gossiping protocols.

$$p_e = p + (q-p)\left(1 - \frac{\min(s_\alpha, t_\alpha)}{t_\alpha}\right) \tag{1}$$

Variants $A-C$ boost p_e for packets early in attempted delivery to combat *early fail*. It would be possible to employ a compound heuristic using more than one of $A-C$, but this would be unlikely to yield much additional improvement as each addresses a similar problem.

Variant A is similar to *GeRaF* [28], and boosts p_e on the *geometric distance from the source location to the current packet location*. The measured packet attribute s_A is simply the Euclidean distance from the source node, at which the packet was generated, to the current node, at which the packet currently resides. This implicitly requires that each node be aware of its geographical position, and that the position of the source node is either known to each candidate relay node, or is encoded into the packet itself. We take the threshold value $t_A = 750$m.

Variant B is similar to *GOSSIP1* [12] and *PPSNRP* [4], boosting p_e on the *number of hops traversed from the source node to the current node*. The measured packet attribute s_B is the number of node-to-node hops encountered by the packet in reaching its current location, either extracted from either a route trace or a simple hop counter encoded into the packet. We take the threshold value $t_B = 3$ hops.

Variant C boosts p_e on the *time elapsed since delivery attempts began*. This is a novel approach not previously considered in the literature to the best of our knowledge, and is anticipated to be of greatest relevance to applications with real-time packet delivery requirements. The measured packet attribute s_B is the wall time that has elapsed since the packet was first made available for broadcast at the source node, until the time at which the probabilistic retransmission decision is made. This implicitly assumes that nodes have access to timing data of sufficient accuracy to support the retransmission decision to the desired level of accuracy. We take the threshold value $t_C = 0.03s$.

Variant D boosts p_e for packets late in attempted delivery to combat *late fail* where significant resources have already been invested in lost packets, an approach similar to *PPSNRP* [4] but replacing network distance with geometric distance similar to *GeRaF* [28]. Variant D boosts p_e on the *geometric distance from the current packet location to the destination location*. The measured packet attribute s_D is simply the Euclidean distance from the current node, at which the packet currently resides, to the destination node, to which the packet is travelling. This implicitly requires that each node be aware of its geographical position, and that the position of the destination node is either known to each candidate relay node, or is encoded into the packet itself. We take the threshold value $t_D = 500m$

In all cases the shapes of curves for variants $A-E$ were similar to those of U, suggesting that the pattern of network behaviour was qualitatively unchanged. However, variants $A-E$ generally showed improved behaviour in some metrics. In figure 2(b) there is little difference in any of the plots, suggesting that the gossip variants are more able to reduce *early fail* effects than *late fail*.

In figures 1(b), 3(b), and to a lesser extent 4(b), we see variant E performs somewhat better than the unmodified protocol U, and that variants $A-D$ perform somewhat better than E but broadly comparable with other variants $A-D$. This suggests some success in combating the *early fail* problem. However, figure 5(b) shows little variation between any plots, and little success in reducing the relative frequency for low-hop delivery attempts for undelivered packets to the low frequencies for successful delivery attempts illustrated in figure 5(a).

We therefore conclude that variants $A-E$ have some small influence in reducing the *early fail* problem, but insufficient influence to solve the problem completely. Further improvement may require a fundamentally different approach. For example, the sensornet designer may select protocols of greater complexity which are at least partially stateful. However, the previous observation that a resource constrained sensornet may be unable to support heavyweight stateful protocols remains relevant and valid.

4.3 Quantitative performance analysis

We now assess the relative merit of variants A-E against unmodified gossip U quantitatively. Metrics $M_1 - M_4$, described below, are shown normalised against the value obtained for the unmodified gossip protocol configuration U in table 1. These metrics pertain to all network behaviour throughout network lifetime, and as such are suitable for analysing relative performance but are not suitable as online heuristics.

M_1: *Latency per metre:* Mean time for a packet to travel 1 metre. Measured in $m^{-1}s$. Defined in the range $(0,\infty)$. Ideal value $M_1 = 0$.

M_2: *Packet delivery failure ratio:* Proportion of packets created at source nodes by the simulated application which the network attempted to deliver, but were lost before reaching their intended destination. Unitless. Defined in the range $[0,1]$. Ideal value $M_2 = 0$.

M_3: *Energy per packet per metre:* Mean energy for 1 packet to travel 1 metre. Measured in $Jpacket^{-1}m^{-1}$. Defined in the range $(0,\infty)$. Ideal value $M_3 = 0$.

M_4: *Delivery path straightness:* Ratio x/y of physical Euclidean distance, x, and actual hop-by-hop delivery route length, y, between source and destination. Unitless. Defined in the range $(0,1]$. Ideal value $M_4 = 1$.

The values displayed in table 1 are unitless as they indicate the relative magnitude in comparison to the unmodified gossip protocol, rather than indicating the measured value itself. All figures are given to three decimal places.

For some metric M_n, a measured value for some protocol variant α is given by $M_{n\alpha}$. For each $n \in \{1,2,3,4\}$ we measure metrics $M_1 - M_4$ using the original gossiping protocol U giving measurements $M_{1U} - M_{4U}$. For each gossiping variant $x \in \{A,B,C,D,E\}$ we similarly measure $M_1 - M_4$ giving measurements $M_{1x} - M_{4x}$. We then compare the values obtained from the original gossiping protocol against each gossiping variant, using the values in 1 showing each x-value as a ratio of the U-value for each metric.

Where $M_{nx} < M_{nU}$ the variant x reduced the metric n, where $M_{nx} > M_{nU}$ the variant x increased the metric, and where $M_{nx} = M_{nU}$ the variant x had no effect on the metric. For metrics $M_1 - M_3$ lower values $M_{n\alpha}$ are desirable, whereas for M_4 higher values are desirable.

Looking at table 1, it is immediately obvious that there is very little variation across the protocol variants in any of the metrics. This supports the findings of section 4.2 in which very little difference in the respective approximated probability distribution functions was observed. The simple context-sensitive heuristic-driven variants of the gossip

Metric	U	A	B	C	D	E
M_1	*1.000*	0.975	0.962	0.957	0.962	0.987
M_2	*1.000*	0.991	1.000	1.002	1.009	0.993
M_3	*1.000*	1.012	1.000	1.014	0.999	1.010
M_4	*1.000*	1.018	1.026	1.035	1.042	1.001

Table 1. Normalised metrics for all variants

protocol considered in section 4 do not perform appreciably better than the simpler, plain gossip protocol.

The variants also do not appear to harm network performance. If a given variant were found to confer some material advantage for some specialised traffic flow type in a given deployment context, it would seem entirely feasible to use this variant with reasonable confidence that more general traffic flows would not be affected adversely. However, this seems insufficient evidence of the merit of such a variant in the general case.

In the absence of demonstrable and meaningful performance improvement we conclude that the additional complexity of the variant protocols cannot be justified, either in terms of the additional computational complexity at sparingly-resourced nodes in executing these protocols or in terms of analytical effort and design understandability.

An alternative approach might simply be to send multiple copies. If interpacket delay is sufficient to guarantee independence of the progress of n send attempts, and the probability of a given attempt succeeding is q, the probability of both attempts failing is $(1-q)^n$ which diminishes quickly as the nth power. If approximately 92% of delivery attempts succeed (see section 4.1) then just two send attempts reduces the probability of both failing to $(1-0.92)^2 = 6.04 \times 10^{-3}$. However, this relies on the probability of successful delivery attempts being entirely independent. This assumption may not hold in sensornets. For example, network congestion induced by an initial attempt may render subsequent attempts less likely to succeed.

5 Conclusions

This paper examines the issue of selecting or constructing protocols with behaviours appropriate to the requirements of a sensornet deployed into a given physical environment. Tuning existing protocols may deliver improved network performance, but without necessarily offering any insight into the underlying issues which may induce substandard performance. Constructing improved protocols for a specific application or a specific sensornet has the potential to deliver optimal performance, but success is unlikely unless these modifications address real problems.

We measured the performance of the gossiping protocol in a repeatable sensornet scenario, then measured the performance of several gossiping protocol variants in the same scenario. It was shown that these variants were sometimes able to deliver small improvements in one or more metrics of interest, but none offered fundamentally different outcomes despite the range of behaviours employed.

It is possible that greater protocol differentiation would be observed if a different sensornet scenario was employed, but the scenario was identical for all protocols and was not designed with any individual protocol in mind.

Several undesirable behaviours relating to undelivered traffic were observed and measured. A number of areas for improvement were identified. In particular, the requirement for online per-packet adaptation is particularly acute at the beginning and end of attempted delivery paths. Future protocol design can address these specific undesirable observed behaviours, rather than addressing perceived problems that may, or may not, or not arise in practice.

The observations and results presented in this paper relate only to those protocols explicitly identified, and do not necessarily hold true for other protocols. However, there is no evidence to suggest that similar trends do not exist for other comparable examples or classes of sensornet protocol.

6 Acknowledgements

This work is partially supported by an EPSRC grant, and by BAE Systems plc under the grant "Hierarchical System Management for Integrated Modular Systems".

References

[1] G. Anastasi, E. Ancillotti, M. Conti, and A. Passarella. Design and performance evaluation of a transport protocol for ad hoc networks. *The Computer Journal*, 52(2):186–209, March 2009.

[2] A. Baker, S. Dixon, F. Drabble, J. Gibbings, A. Lewkowicz, D. Moffat, and R. Shaw. *The Systematic Experiment*. Cambridge University Press, Cambridge, 1986.

[3] G. Barrenetxea, F. Ingelrest, G. Schaefer, and M. Vetterli. The hitchhiker's guide to successful wireless sensor network deployments. In *Proceedings of the 6th ACM conference on Embedded Network Sensor Systems*, pages 43–56, Raleigh, NC, 5-7 November 2008. ACM Press, New York, NY.

[4] C. Barrett, S. Eidenbenz, L. Kroc, M. Marathe, and J. Smith. Parametric probabilistic routing in sensor networks. *Mobile Networks and Applications*, 10(4):529–544, 2005.

[5] G. Box, J. Hunter, and W. Hunter. *Statistics for Experimenters*. Wiley, Hoboken, NJ, 2nd edition, 2005.

[6] M. Ceriotti, L. Mottola, G. Picco, A. Murphy, S. Gună, M. Corrà, M. Pozzi, D. Zonta, and P. Zanon. Monitoring heritage buildings with wireless sensor networks: the Torre Aquila deployment. In *Proceedings of the 8th International Conference on Information Processing in Sensor Networks*,

pages 277–288, San Francisco, CA, 13-16 April 2009. IEEE Computer Society, Los Alamitos, CA.

[7] Crossbow Technology Inc. MICA2 datasheet, part number 6020-0042-08 Rev A. http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/MICA2_Datasheet.pdf. Accessed 09/01/2008.

[8] J. Doyle. *Routing TCP/IP*, volume 1. Cisco Press, Indianapolis, IN, 1998.

[9] C. Fullmer and J. Garcia-Luna-Aceves. Solutions to hidden terminal problems in wireless networks. In *Proceedings of the 22nd ACM SIGCOMM conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, pages 39–49, Cannes, 16-18 September 1997. ACM Press, New York, NY.

[10] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker. Complex behavior at scale: An experimental study of low-power wireless sensor networks. Technical Report CSD-TR 02-0013, UCLA, Los Angeles, CA, February 2002.

[11] M. Grenier and N. Navet. Fine-tuning MAC-level protocols for optimized real-time QoS. *IEEE Transactions on Industrial Informatics*, 4(1):6–15, February 2008.

[12] Z. Haas, J. Halpern, and L. Li. Gossip-based ad hoc routing. *IEEE Transactions on Networking*, 14(3):479–491, 2006.

[13] T. He, B. Blum, Q. Cao, J. Stankovic, S. Son, and T. Abdelzaher. Robust and timely communication over highly dynamic sensor networks. *Real-Time Systems*, 37(3):261–289, 2007.

[14] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister. System architecture directions for networked sensors. In *Proceedings of the 5th International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 93–103, Cambridge, MA, 7-11 March 2000. ACM Press, New York, NY.

[15] X. Jiang, J. Taneja, J. Ortiz, A. Tavakoli, P. Dutta, J. Jeong, D. Culler, P. Levis, and S. Shenker. An architecture for energy management in wireless sensor networks. *ACM SIGBED Review*, 4(3):31–36, 2007.

[16] J. Kulik, W. Heinzelman, and H. Balakrishnan. Negotiation-based protocols for disseminating information in wireless sensor networks. *Wireless Networks*, 8(2/3):169–185, 2002.

[17] X. Li, K. Moaveninejad, and O. Frieder. Regional gossip routing for wireless ad hoc networks. *Mobile Network Applications*, 10(1-2):61–77, 2005.

[18] A. Mohan, H. Wei, D. Gay, P. Buonadonna, and A. Mainwaring. End-to-end performance characterization of sensornet multi-hop routing. In *Proceedings of the 2nd IEEE International Conference on Pervasive Services*, pages 27–36, Santorini, 11-14 July 2005. IEEE Computer Society, Los Alamitos, CA.

[19] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proceedings of the 5th ACM/IEEE International Conference on Mobile Computing and Networking*, pages 151–162, Seattle, WA, 15-19 August 1999. ACM Press, New York, NY.

[20] V. Park and M. Corson. A highly adaptive distributed routing algorithm for mobile wireless networks. In *Proc. 16th IEEE International Conference on Computer Communications*, pages 1405–1413, Las Vegas, NV, 22-25 September 1997. IEEE Computer Society, Los Alamitos, CA.

[21] J. Polastre, J. Hui, P. Levis, J. Zhao, D. Culler, S. Shenker, and I. Stoica. A unifying link abstraction for wireless sensor networks. In *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, pages 76–89, San Diego, CA, 2-4 November 2005. ACM Press, New York, NY.

[22] V. Raghunathan, C. Schurgers, S. Park, and M. Srivastava. Energy aware wireless microsensor networks. *IEEE Signal Processing Magazine*, 19(2):40–50, March 2002.

[23] T. Roosta, M. Menzo, and S. Sastry. Probabilistic geographic routing in ad hoc and sensor networks. In *Proceedings of the 2nd International Workshop on Wireless Ad-hoc Networks*, pages 1–9, London, 23-26 May 2005.

[24] J. Tate and I. Bate. YASS: A scaleable sensornet simulator for large scale experimentation. In *Proceedings of the 31st Communicating Process Architectures Conference*, pages 411–430, York, 7-10 September 2008. IOS Press, Amsterdam.

[25] J. Tate and I. Bate. Sensornet protocol tuning using principled engineering methods. *The Computer Journal*, 2009. Advance Access publication on 19 August 2009 at http://comjnl.oxfordjournals.org/cgi/content/abstract/bxp077.

[26] J. Tate and I. Bate. Understanding behavioural tradeoffs in large-scale sensornet design. In *Proceedings of the 1st IEEE International Workshop on Quantitative Evaluation of Large-Scale Systems and Technologies*, pages 1085–1091, Bradford, 26-29 May 2009. IEEE Press, New York, NY.

[27] M. Totaro and D. Perkins. Using statistical design of experiments for analyzing mobile ad hoc networks. In *Proceedings of the 8th International Symposium on Modeling Analysis and Simulation of Wireless and Mobile Systems*, pages 159–168, Montreal, 10-13 October 2005. ACM Press, New York, NY.

[28] M. Zorzi and R. Rao. Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance. *IEEE Transactions on Mobile Computing*, 2(4):337–348, 2003.