Networking Working Group
Internet-Draft
Intended status: Informational
Expires: August 18, 2009

P. Levis
Stanford University
A. Tavakoli
S. Dawson-Haggerty
UC Berkeley

Feb 14, 2009

Overview of Existing Routing Protocols for Low Power and Lossy Networks

draft-ietf-roll-protocols-survey-06

Status of this Memo

Copyright Notice

Abstract

Low-power wireless devices, such as sensors, actuators and smart objects, present difficult constraints: very limited memory, little processing power, and long sleep periods. As most of these devices are battery-powered, energy efficiency is critically important. Wireless link qualities can vary significantly over time, requiring protocols to make agile decisions yet minimize topology change energy costs. Routing over such low power and lossy networks introduces requirements that existing routing protocols may not fully address. Using existing application requirements documents, this document derives a minimal and not exhaustive set of criteria for routing in low-power and lossy networks. It provides a brief survey of the strengths and weaknesses of existing protocols with respect to these criteria. From this survey it examines whether existing and mature IETF protocols can be used without modification in these networks, or whether further work is necessary. It concludes that no existing IETF protocol meets the requirements of this domain.

Table of Contents

1. Terminology

AODV: Ad-hoc On Demand Vector Routing

DSR: Dynamic Source Routing

DYMO: Dynamic Mobile On-Demand

IS-IS: Intermediate System to Intermediate System

OLSR: Optimized Link State Routing

OSPF: Open Shortest Path First

RIP: Routing Information Protocol

TBRPF: Topology Dissemination Based on Reverse Path Forwarding

LLN: Low power and Lossy Network

LSA: Link State Advertisement

LSDB: Link State Database

MANET: Mobile Ad-hoc Network

MAC: Medium Access Control

MPLS: Multiprotocol Label Switching

MPR: Multipoint Relays

MTU: Maximum Transmission Unit

ROLL: Routing in Low power and Lossy Networks

TDMA: Time Division Multiple Access

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Introduction

Wireless is increasingly important to computer networking. As the technological progress behind Moore's Law has reduced computer prices

and form factors, networking has come to include not only servers and desktops, but laptops, palmtops, and cellphones. As computing device costs and sizes have shrunk, small wireless sensors, actuators, and smart objects have emerged as an important next step. The sheer number of such low-power networked devices means that they cannot depend on human intervention (e.g., adjusting position) for good connectivity: they must have routing protocols that enable them to self-organize into multihop networks.

Energy is a fundamental challenge in these devices. Convenience and ease of use requires they be wireless and therefore battery powered. Low power operation is a key concern for these sensors and actuators so as to allow them to function for months and years without interruption. Cost points and energy limitations cause these devices to have very limited computational and storage resources: a few kB of RAM and a few MHz of CPU is typical. As energy efficiency does not improve with Moore's Law, these limitations are not temporary. This trend towards smaller, lower power, and more numerous devices has led to new low-power wireless link layers to support them.

In practice, wireless networks observe much higher loss rates than wired ones do, and low-power wireless is no exception. Furthermore, many of these networks will include powered as well as energy constrained nodes. Nevertheless, for cost and scaling reasons, many of these powered devices will still have limited resources.

These low power and lossy networks introduce constraints and requirements that other networks typically do not possess; for instance, in addition to the constraints of limited resources and small power sources which constrain the amount of traffic a protocol may generate, these applications demand an embrace of heterogeneous node capabilities, and good support for specific traffic patterns ([I-D.ietf-roll-home-routing-reqs] and [I-D.ietf-roll-indus-routing-reqs]).

2.1. LLN Properties

In short, LLN routing protocols must operate under a set of constraints that traditional protocols typically do not consider.

- They need to operate with a hard, very small bound on state.

- In most cases, they optimize for preserving energy.

- Typical application data patterns are not simply unicast flows.

- They need to be effective with very small packet sizes, such as less than 127 octets.

- They have to be very careful when trading off in-node efficiency for generality; most LLN nodes don't have computational or memory resources to waste.

2.2. Question to Answer

As existing protocols were not designed with all of these constraints in mind, they have made trade-offs which may or may not be appropriate for LLNs. The first step to reaching consensus on a routing protocol for LLNs is to decide which of these two is true. If an existing protocol can meet LLN requirements without any changes, then barring extenuating circumstances, it behooves us to use such an existing protocol. However, if no current protocol can meet LLN's requirements, then further work will be needed to define and standardize a protocol that can. Whether or not such a protocol involves extensions to an existing protocol or developing a new protocol is outside the scope of this document: this document simply seeks to answer the question: do LLNs require a new protocol specification document at all?

3. Methodology

To answer the question of whether LLNs require new protocol specification work, this document examines existing routing protocols and how well they can be applied to low power and lossy networks. It provides a set of criteria with which to compare the costs and benefits of different protocol designs and examines existing protocols in terms of these criteria.

3.1. Protocols Considered

This document does not seek to answer the question of whether there is any protocol outside the IETF which could meet LLN application requirements. Rather, it seeks to answer whether any existing protocols developed and specified by the IETF can meet such requirements. If an existing protocol specification can be used unchanged, then writing additional protocol specifications is unnecessary. There are many academic papers and experimental protocol implementations available. While one or more of these may meet LLN requirements, if they are not specified in an RFC then a working group will need to specify those in a new RFC for them to be a standard. The question this document seeks to answer is not whether proposed, evaluated, theoretical or hypothetical protocol designs can satisfy LLN requirements: the question is whether existing IETF protocols can.

Therefore, this document considers "existing routing protocols" to be

protocols that are specified in RFCs or, in the cases of DYMO [I-D.ietf-manet-dymo] or OLSRv2 [I-D.ietf-manet-olsrv2], a very stable and mature draft that is a charter item of an active IETF working group. The list of considered protocols is OSPF [RFC2328], IS-IS [RFC1142], RIP [RFC2453], OLSR [RFC3626], OLSv2 [I-D.ietf-manet-olsrv2], TBRPF [RFC3684], AODV [RFC3561], DYMO [I-D.ietf-manet-dymo], and DSR [RFC4728]. This document also considers notable variants of these protocols, such as Triggered RIP [RFC2091].

Some of these protocols are still works in progress, and so are changing over time. To enable this document to be correct yet not dependent on this evolution, it document considers specifications as of a specific date: November 31, 2008.

This document does not consider DTN bundles [RFC5050] or the DTN Licklider protocol [RFC5326] as suggested by the ROLL working group charter, because they are not routing protocols. This document does not consider the DTN routing protocol PRoPHET [I-D.irtf-dtnrg-prophet] because its design is based on the non-randomness of node mobility, which is not common to LLN application domains.

3.2. Criteria

The five criteria this document uses are derived from a set of documents that describe the requirements of a few major LLN application scenarios. The five criteria, presented in Section 4, are neither exhaustive nor complete. Instead, they are one specific subset of high-level requirements shared across all of the application requirement drafts. Because every application requirement draft specifies these criteria, then a protocol which does not meet one of them cannot be used without modifications or extensions. However, because these criteria represent a subset of the intersection of the application requirements, any given application domain may impose additional requirements which a particular protocol may not meet. For this reason, these criteria are "necessary but not sufficient." A protocol that does not meet the criteria cannot be used as specified, but it is possible that a protocol meets the criteria yet is not able to meet the requirements of a particular application domain. Nevertheless, a protocol that meets all of the criteria would be very promising, and deserve a closer look and consideration in light of LLN application domains.

3.3. Evaluation

This document evaluates the above protocols by thinking through each specification and considering a hypothetical implementation that

performs as well as possible on the criteria. The evaluation is based on what a specification allows, rather than any particular implementation of that specification. For example, while many DYMO implementations use hopcount as a routing metric, the DYMO specification allows a hop to add more than one to the routing metric, so DYMO as a specification can support some links or nodes being more costly than others. The analysis does not consider hypothetical extensions to protocols that require additional fields or message exchanges.

4. Criteria

This section presents five important criteria for routing in low power and lossy networks. Later sections evaluate protocols against them. The evaluation attempts to take a complicated and interrelated set of design decisions and trade-offs and condense them to a simple "pass", "fail", or "?". As with any simplification, there is a risk of removing some necessary nuance. However, we believe that being forced to take a position on whether or not these protocols are acceptable according to binary criteria is constructive.

We derive these criteria from existing documents that describe ROLL network application requirements [I-D.ietf-roll-home-routing-reqs] [I-D.ietf-roll-urban-routing-reqs] [I-D.ietf-roll-indus-routing-reqs]. These criteria do not encompass all application requirements. Instead, they are a common set of routing protocol requirements that the applications domains in these documents share. Considering this very general and common set of requirements sets a minimal bar for a protocol to be applicable for an LLN deployment.

If a protocol cannot meet these minimalist criteria, then it cannot be used unchanged in several major ROLL application domains and so is unlikely to be a good candidate for use within the broader scope of all LLN application domains. Satisfying these minimal criteria is necessary but not sufficient. They do not represent the complete intersection of application requirements and applications introduce additional, more stringent requirements. But this simplified view provides a first cut of the applicability of existing protocols, and those that do satisfy them might be reasonable candidates for further study.

The five criteria are "routing state", "loss response", "control cost", "link cost", and "node cost". For each of these, the value "pass" indicates that a given protocol has satisfactory performance according to the criterion. The value "fail" indicates that the protocol does not have acceptable performance according to the

criterion, and that the document defining the protocol does not, as written, contain sufficient flexibility to allow the protocol to meet the criterion while conforming to the specification. Finally, "?" indicates that an implementation could exhibit satisfactory performance while following the protocol specification, but that the implementation decisions necessary to do so are not specified and may require some exploration. In other words, a "fail" means a protocol would have to be modified so it is not compliant with its specification in order to meet the criterion, while a "?" means a protocol would require a supplementary document further constraining and specifying how a protocol should behave.

A protocol failing to meet one or more of the criteria does not exclude it from being used in LLNs. Rather, it means that, in order to be used in LLNs, the protocol would need to be extended in ways that do not conform with the current specification document.

4.1. Formal Definitions

To provide precise definitions of these criteria, we use formal big-O notation, where N refers to the number of nodes in the network, D refers to the number of unique destinations, and L refers to the size of a node's local, single-hop neighborhood (the network density). We explain the derivation of each criterion from application requirements in its corresponding section.

4.2. Routing State

This criterion indicates whether routing state scales reasonably within the memory resources of low-power nodes. According to this criterion, routing state that scales linearly with the size of the network or a node's neighborhood fail. Scaling with the size of the network prevents networks from growing to to the sizes necessary for many LLN applications when faced with the memory constraints devices in such applications exhibit. Similarly, scaling with the network density precludes dense deployments.

However, as many low-power and lossy networks behave principally as data collection networks and principally communicate through routers to data collection points in the larger Internet, scaling with the number of such collection points is reasonable. Protocols whose state scales with the number of destinations pass.

More precisely, routing state scaling with O(N) or O(L) fails. State that scales O(D) (assuming no N or L) passes.

4.3. Loss Response

In low power and lossy networks, links routinely come and go due to being close to the signal-to-noise threshold at the physical layer. It is important that link churn not trigger unnecessary responses by the routing protocol. This point is stressed in all the application requirement documents, pointing to the need to localize response to link failures with no triggering of global network re-optimization, whether for reducing traffic or for maintaining low route convergence times ([I-D.ietf-roll-home-routing-reqs], [I-D.ietf-roll-urban-routing-reqs], and [I-D.ietf-roll-indus-routing-reqs]). The industrial routing requirements draft states that protocols must be able to "recompute paths based on underlying link characteristics which may change dynamically", as well as reoptimize when the device set changes to maintain service requirements. The protocol should also "always be in the process of optimizing the system in response to changing link statistics." Protocols with these properties should take care not to require global updates.

A protocol which requires many link changes to propagate across the entire network fails. Protocols which constrain the scope of information propagation to only when they affect routes to active destinations, or to local neighborhoods, pass. Protocols which allow proactively path maintenance pass if the choice of which paths to maintain is user-specified.

More precisely, loss responses that require O(N) transmissions fail, while responses that can rely on O(1) local broadcasts or O(D) route updates pass.

4.4. Control Cost

Battery-operated devices are a critical component of all three application spectrums, and as such special emphasis is placed on minimizing power consumption to achieve long battery lifetime, [I-D.ietf-roll-home-routing-reqs], with multi-year deployments being a common case [I-D.ietf-roll-indus-routing-reqs]. In terms of routing structure, any proposed LLN routing protocol ought to support the autonomous organization and configuration of the network at the lowest possible energy cost [I-D.ietf-roll-urban-routing-reqs].

All routing protocols must transmit additional data to detect neighbors, build routes, transmit routing tables, or otherwise conduct routing. As low-power wireless networks can have very low data rates, protocols which require a minimum control packet rate can have an unbounded control overhead per data packet. This is particularly true for event-driven networks, which only report data

when certain conditions are met. Regions of a network which never meet the condition can be forced to send significant control traffic even when there is no data to send. For these use cases, hard-coded timing constants are unacceptable, because they imply a prior knowledge of the expected data rate.

Of course, protocols require the ability to send at least a very small amount of control traffic, in order to discover a topology. But this bootstrapping discovery and maintenance traffic should be small: communicating once an hour is far more reasonable than communicating once a second. So while control traffic should be bounded by data traffic, it requires some leeway to bootstrap and maintain a long-lived yet idle network.

In the case of control traffic, the communication rate (sum of transmissions and receptions at a node) is a better measure than the transmission rate (since energy is consumed for both transmissions and receptions). Controlling the transmission rate is insufficient, as it would mean that the energy cost (sum of transmission and receptions) of control traffic could grow with O(L).

A protocol fails the control cost criterion if its per-node control traffic (transmissions plus receptions) rate is not bounded by the data rate plus a small constant. For example, a protocol using a beacon rate only passes if it can be turned arbitrarily low, in order to match the data rate. Furthermore, packet losses necessitate that the control traffic may scale within a O(log(L)) factor of the data rate. Meaning, if R is the data rate and e is the small constant, then a protocol's control traffic must be on the order of O(R log(L) + e) to pass this criteria. The details of why O(log(L)) is necessary are in Appendix B.

4.5. Link and Node Cost

These two criteria specify how a protocol chooses routes for data packets to take through the network. Classical routing algorithms typically acknowledge the differing costs of paths and may use a shortest path algorithm to find paths. This is a requirement for low power networks, as links must be evaluated as part of an objective function across various metric types, such as minimizing latency and maximizing reliability [I-D.ietf-roll-indus-routing-reqs].

However, in low power networks it is also desirable to account for the cost of forwarding through particular routers. Applications require node or parameter constrained routing, which takes into account node properties and attributes such as power, memory, and battery life that dictate a router's willingness or ability to route other packets. Home routing requirements note that devices will vary

in their duty cycle, and that routing protocols should prefer nodes with permanent power [I-D.ietf-roll-home-routing-reqs]. The urban requirements note that routing protocols may wish to take advantage of differing data processing and management capabilities among network devices [I-D.ietf-roll-urban-routing-reqs]. Finally, industrial requirements cite differing lifetime requirements as an important factor to account for [I-D.ietf-roll-indus-routing-reqs]. Node cost refers to the ability for a protocol to incorporate router properties into routing metrics and use node attributes for constraint-based routing.

A "pass" indicates that the protocol contains a mechanism allowing these considerations to be considered when choosing routes.

5. Routing Protocol Taxonomy

Routing protocols broadly fall into two classes: link-state and distance-vector.

A router running a link-state protocol first establishes adjacency with its neighbors and then reliably floods the local topology information in the form of a Link State Advertisement packet. The collection of LSAs constitutes the Link State Database (LSDB) that represents the network topology, and routers synchronize their LSDBs. Thus each node in the network has a complete view of the network topology. Each router uses its LSDB to compute a routing table where each entry (reachable IP destination address) points to the next hop along the shortest path according to some metric. Link state protocols (such as OSPF and IS-IS) support the concept of area (called "level" for IS-IS) whereby all the routers in the same area share the same view (they have the same LSDB) and areas are interconnected by border routers according to specific rules that advertise IP prefix reachability between areas.

A distance vector protocol exchanges routing information rather than topological information. A router running a distance vector protocol exchanges information with its "neighbors" with which it has link layer connectivity. Tunneling and similar mechanisms can virtualize link layer connectivity to allow neighbors that are multiple layer 2 hops away. Rather than a map of the network topology from which each router can calculate routes, a distance vector protocol node has information on what routes its neighbors have. Each node's set of available routes is the union of its neighbors routes plus a route to itself. In a distance vector protocol, nodes may only advertise routes which are in use, enabling on-demand discovery. In comparison to link state protocols, distance vector protocols have the advantage of only requiring neighbor routing information, but also have

corresponding limitations which protocols must address, such as routing loops, count to infinity, split horizon, and slow convergence times. Furthermore, routing constraints are difficult to enforce with distance vector protocols.

Neighbor discovery is a critical component of any routing protocol. It enables a protocol to learn about which other nodes are nearby and which it can use as the next hop for routes. As neighbor discovery is a key component of many protocols, several general protocols and protocol mechanisms have been designed to support it. A protocol's neighbor set is defined by how many "hops" away the set reaches. For example, the 1-hop neighbor set of a node is all nodes it can directly communicate with at the link layer, while the 2-hop neighbor set is its own 1-hop neighbor set and the 1-hop neighbor sets of all of its 1-hop neighbors.

Because nodes often have very limited resources for storing routing state, protocols cannot assume that they can store complete neighbor information. For example, a node with 4kB of RAM, a typical amount for top-end microcontrollers, cannot store full neighbor state when it has 1000 other nodes nearby. This means that ROLL protocols must have mechanisms to decide which of many possible neighbors they monitor as routable next hops. For elements such as 2-hop neighborhoods, these decisions can have a significant impact on the topology that other nodes observe, and therefore may require intelligent logic to prevent effects such as network partitions.

5.1. Protocols Today

Wired networks draw from both approaches. OSPF or IS-IS, for example, are link-state protocols, while RIP is a distance-vector protocol.

MANETs similarly draw from both approaches. OLSR is a link-state protocol, while AODV and DYMO are distance vector protocols. The general consensus in core networks is to use link state routing protocols as IGPs for a number of reasons: in many cases having a complete network topology view is required to adequately compute the shortest path according to some metrics. For some applications such as MPLS Traffic Engineering it is even required to have the knowledge of the Traffic Engineering Database for constraint based routing.

Furthermore link state protocols typically have superior convergence speeds (ability to find an alternate path in case of network element failure), are easier to debug and troubleshoot, and introduce less control packet overhead than distance vector protocols. In contrast, distance vector protocols are simpler, require less computation, and have smaller storage requirements. Most of these tradeoffs are

similar in wireless networks, with one exception. Because wireless links can suffer from significant temporal variation, link state protocols can have higher traffic loads as topology changes must propagate globally, while in a distance vector protocol a node can make local routing decisions with no effect on the global routing topology.

One protocol, DSR, does not easily fit into one of these two classes. Although it is a distance vector protocol, DSR has several properties that make it differ from most other protocols in this class. We examine these differences in our discussion of DSR.

The next two sections summarize several well established routing protocols. The table below shows, based on the criteria described above, whether these protocols meet ROLL criteria. Appendix A contains the reasoning behind each value in the table.

Protocol	State	Loss	Control	Link Cost	Node Cost
OSPF/IS-IS	fail	fail	fail	pass	fail
OLSRv2	fail	?	?	pass	pass
TBRPF	fail	pass	fail	pass	?
RIP	pass	fail	pass	?	fail
AODV	pass	fail	pass	fail	fail
DYMO	pass	?	pass	?	fail
DSR	fail	pass	pass	fail	fail

Figure 1

6. Link State Protocols

6.1. OSPF & IS-IS

OSPF (specified in [RFC2328] for IPv4 and in [RFC2740] for IPv6)) is a link state protocol designed for routing within an Internet Autonomous System (AS). OSPF provides the ability to divide a network into areas, which can establish a routing hierarchy. The topology within an area is hidden from other areas and IP prefix reachability across areas (inter-area routing) is provided using summary LSAs. The hierarchy implies that there is a top-level routing area (the backbone area) which connects other areas. Areas may be connected to the back-bone area through a virtual link. OSPF maintains routing adjacencies by sending hello messages. OSPF calculates the shortest path to a node using link metrics (that may reflect the link bandwidth, propagation delay, ...). OSPF Traffic Engineering (OSPF-TE, [RFC3630]) extends OSPF to include information

on reservable, unreserved, and available bandwidth.

IS-IS (specified in [RFC1142]) is similar in many respects to OSPF, but as a descendent of the OSI protocol suite differs in some places such as the way areas are defined and used. However, routing adjacencies are also maintained by local propagation of the LSDB, and a shortest path computation is used over a metric space which may measure delay, errors, or other link properties.

6.2. OLSR & OLSRv2

Optimized Link State Routing (OLSR) (see [RFC3626] and [I-D.ietf-manet-olsrv2]) is a link state routing protocol for MANETs. OLSR routers flood link state advertisement packets throughout the entire network, such that each node has a map of the network topology. Because link variations can lead to heavy flooding traffic when using a link state approach, OLSR establishes a topology for minimizing this communication, imposes minimum time interval between two successive control transmissions by a router, and makes triggered updates optional. Each node maintains a set of nodes called its Multipoint Relays (MPR), which is a subset of the one-hop neighbors whose connectivity covers the two-hop neighborhood. Each node that is an MPR maintains a set called its MPR selectors, which are nodes that have chosen it to be an MPR.

OLSR uses these two sets to apply three optimizations. First, only MPRs generate link state information. Second, nodes use MPRs to limit the set of nodes that forward link state packets. Third, an MPR, rather than advertise all of its links, can advertise only links to its MPR selectors. Together, these three optimizations can greatly reduce the control traffic in dense networks, as the number of MPRs should not increase significantly as a network becomes denser.

OLSR selects routes based on hop counts, and assumes an underlying protocol that determines whether a link exists between two nodes. OLSR's optimized flooding allows it to quickly adapt to and propagate topology changes.

OLSR is closely related to clustering algorithms in the wireless sensor networking literature, in which cluster heads are elected such that routing occurs over links between cluster heads and all other nodes are leaves that communicate to a cluster head.

6.3. TBRPF

Topology Dissemination Based on Reverse Path Forwarding (see [RFC3684]) is another proactive link state protocol for MANETs.

TBRPF computes a source tree, which provides routes to all reachable nodes. It reduces control packet overhead by having nodes only transmit a subset of their source tree as well as by using differential updates.

The major difference between TBRPF and OLSR is the routing data that nodes advertise and who chooses to aggregate information. In OLSR, nodes select neighbors to be MPRs and advertise their link state for them; in TBRPF, nodes elect themselves to advertise relevant link state based on whether it acts as a next hop.

7. Distance Vector protocols

7.1. RIP

The Routing Information Protocol (RIP) (defined in [RFC2453]) predates OSPF. As it is a distance vector protocol, routing loops can occur and considerable work has been done to accelerate convergence since the initial RIP protocols were introduced. RIP measures route cost in terms of hops, and detects routing loops by observing a route cost approach infinity where "infinity" is referred to as a maximum number of hops. RIP is typically not appropriate for situations where routes need to be chosen based on real-time parameters such as measured delay, reliability, or load or when the network topology needs to be known for route computation.

"Triggered RIP" (defined in [RFC2091]) was originally designed to support "on-demand" circuits. The aim of triggered RIP is to avoid systematically sending the routing database on regular intervals. Instead, triggered RIP sends the database when there is a routing update or a next hop adjacency change: once neighbors have exchanged their routing database, only incremental updates need to be sent. Because incremental updates cannot depend on periodic traffic to overcome loses, triggered RIP uses acknowledgment based mechanisms for reliable delivery.

7.2. Ad-hoc On Demand Vector Routing (AODV)

AODV (specified in [RFC3561]) is a distance vector protocol for MANETs. When one AODV node requires a route to another, it floods a request in the network to discover a route. A depth-scoped flooding process avoids discovery from expanding to the most distant regions of the network that are in the opposite direction of the destination. AODV chooses routes that have the minimum hop count.

If an AODV route request reaches a node that has a route to the destination (this includes the destination itself), that node sends a

reply along the reverse route. All nodes along the reverse route can cache the route. When routes break due to topology changes, AODV floods error messages and issues a new request. Because AODV is on-demand it only maintains routes for active nodes. When a link breaks, AODV issues a Route Error (RERR) and a new route request message (RREQ), with a higher sequence number so nodes do not respond from their route caches. These packets can flood the entire network.

7.3. DYMO

Dynamic Mobile On-Demand routing (DYMO) ([I-D.ietf-manet-dymo]) is an evolution of AODV. The basic functionality is the same, but it has different packet formats, handling rules, and supports path accumulation. Path accumulation allows a single DYMO route request to generate routes to all nodes along the route to that destination. Like AODV, DYMO uses a distance value as its routing metric which must be at least the hop count, but allows DYMO to represent link costs. Like AODV, on link breaks DYMO issues a new route request message (RREQ), with a higher sequence number so nodes do not respond from their route caches. Correspondingly, a route request can flood the entire network.

7.4. DSR

Dynamic Source Routing ([RFC4728]) is a distance vector protocol for MANETs, but a DSR packet source explicitly specifies the route for each packet. Because the route is determined at a single place -- the source -- DSR does not require sequence numbers or other mechanisms to prevent routing loops, as there is no problem of inconsistent routing tables. Unlike AODV and DYMO, by pushing state into packet headers, DSR does not require per-destination routing state. Instead, a node originating packets only needs to store a spanning tree of the part of the network it is communicating with.

8. Neighbor Discovery

A limit on maintained routing state (light footprint) prevents ROLL protocols from assuming they know all 1-hop, 2-hop, or N-hop neighbors. For this reason, while protocols such as MANET-NHDP ([I-D.ietf-manet-nhdp]) and IPv6's neighbor discovery ([RFC4861]) provide basic mechanisms for discovering link-layer neighbors, not all of their features are relevant. This section describes these two protocols, their capabilities, and how ROLL protocols could leverage them.

8.1. IPv6 Neighbor Discovery

IPv6 neighbor discovery provides mechanisms for nodes to discover single-hop neighbors as well as routers that can forward packets past the local neighborhood. There is an implicit assumption that the delegation of whether a node is a router or not is static (e.g., based on a wired topology). The fact that all routers must respond to a Router Solicitation requires that the number of routers with a 1-hop neighborhood is small, or there will be a reply implosion. Furthermore, IPv6 neighbor discovery's support of address autoconfiguration assumes address provisioning, in that addresses reflect the underlying communication topology. IPv6 neighbor discovery does not consider asymmetric links. Nevertheless, it may be possible to extend and adapt IPv6's mechanisms to wireless in order to avoid response storms and implosions.

8.2. MANET-NHDP

The MANET Neighborhood Discovery Protocol (MANET-NHDP) provides mechanisms for discovering a MANET router's symmetric 2-hop neighborhood. It maintains information on discovered links, their interfaces, status, and neighbor sets. MANET-NHDP advertises a node's local link state; by listening to all of its 1-hop neighbor's advertisements, a node can compute its 2-hop neighborhood. MANET-NHDP link state advertisements can include a link quality metric. MANET-NHDP's node information base includes all interface addresses of each 1-hop neighbor: for low-power nodes, this state requirement can be difficult to support.

9. Conclusion

Figure 1 shows that no existing IETF protocol specification meets the criteria described in Section 4. Therefore, having a routing protocol for LLNs requires new protocol specification documents. Whether such documents describe modifications to existing protocols or new protocols it outside the scope of this document and warrants further discussion. However, the results in Figure 1 may provide some insight or guidance in such a discusssion, indicating what protocol mechanisms may be better suited to LLNs than others.

Such a discussion should not, however, be limited to the protocols listed in Figure Figure 1. There are many existing protocols which are unsuitable as a general routing protocol but describe mechanisms that could be very useful in the context of LLNs. Any such future discussion ought to consider how routing in LLNs may benefit from examining mechanisms from a broader suite of protocols than those listed in Figure Figure 1.

10. Security Considerations

LLNs have security considerations. These considerations vary greatly depending on application domain. For example, deployers industrial monitoring networks may impose more stringent confidentiality requirements than home automation networks do. Such requirements are an important consideration in protocol design, but their variety makes distilling them to a minimalist set of "necessary but not sufficient" criteria is of limited use. The criteria in this document are not the only requirements and considerations in LLN protocols, and as such the omission of a security criterion should not be interpreted as a lack of a need for security in LLNs.

11. IANA Considerations

This document includes no request to IANA.

12. Acknowledgements

The authors would like to thank all the members of the ROLL working group for their valuable comments, and the chairs for their helpful guidance.

We are also indebted to the Sensor Network Architecture group at Berkeley for contributing their helpful analysis: Prabal Dutta, Rodrigo Fonseca, Xiaofan Jiang, Jaein Jeong, Jorge Ortiz, and Jay Tanega.

13. References

13.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

13.2. Informative References

[I-D.ietf-manet-dymo]
Chakeres, I. and C. Perkins, "Dynamic MANET On-demand (DYMO) Routing", draft-ietf-manet-dymo-15 (work in progress), November 2008.

[I-D.ietf-manet-nhdp]
Clausen, T., Dearlove, C., and J. Dean, "MANET Neighborhood Discovery Protocol (NHDP)",

draft-ietf-manet-nhdp-07 (work in progress), July 2008.

[I-D.ietf-manet-olsrv2]
Clausen, T., Dearlove, C., and P. Jacquet, "The Optimized Link State Routing Protocol version 2", draft-ietf-manet-olsrv2-07 (work in progress), July 2008.

[I-D.ietf-roll-home-routing-reqs]
Porcu, G., "Home Automation Routing Requirements in Low Power and Lossy Networks", draft-ietf-roll-home-routing-reqs-06 (work in progress), November 2008.

[I-D.ietf-roll-indus-routing-reqs]
Networks, D., Thubert, P., Dwars, S., and T. Phinney, "Industrial Routing Requirements in Low Power and Lossy Networks", draft-ietf-roll-indus-routing-reqs-03 (work in progress), December 2008.

[I-D.ietf-roll-urban-routing-reqs]
Dohler, M., Watteyne, T., Winter, T., Barthel, D., Jacquenet, C., Madhusudan, G., and G. Chegaray, "Urban WSNs Routing Requirements in Low Power and Lossy Networks", draft-ietf-roll-urban-routing-reqs-03 (work in progress), January 2009.

[I-D.irtf-dtnrg-prophet]
Lindgren, A. and A. Doria, "Probabilistic Routing Protocol for Intermittently Connected Networks", draft-irtf-dtnrg-prophet-01 (work in progress), November 2008.

[RFC1142] Oran, D., "OSI IS-IS Intra-domain Routing Protocol", RFC 1142, February 1990.

[RFC2091] Meyer, G. and S. Sherry, "Triggered Extensions to RIP to Support Demand Circuits", RFC 2091, January 1997.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC2453] Malkin, G., "RIP Version 2", STD 56, RFC 2453, November 1998.

[RFC2740] Coltun, R., Ferguson, D., and J. Moy, "OSPF for IPv6", RFC 2740, December 1999.

[RFC3561] Perkins, C., Belding-Royer, E., and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", RFC 3561,

July 2003.

[RFC3626] Clausen, T. and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)", RFC 3626, October 2003.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3684] Ogier, R., Templin, F., and M. Lewis, "Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)", RFC 3684, February 2004.

[RFC4728] Johnson, D., Hu, Y., and D. Maltz, "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4", RFC 4728, February 2007.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC5050] Scott, K. and S. Burleigh, "Bundle Protocol Specification", RFC 5050, November 2007.

[RFC5326] Ramadas, M., Burleigh, S., and S. Farrell, "Licklider Transmission Protocol - Specification", RFC 5326, September 2008.

Appendix A. Routing protocol scalability analysis

This aim of this Appendix is to provide the details for the analysis routing scalability analysis.

"OSPF & IS-IS"

OSPF floods link state through a network. Each router must receive this complete link set. OSPF fails the routing state criterion because it requires each router to discover each link in the network, for a total routing table size which is O(N * L). This also causes it to fail the control cost criterion, since this information must be propagated. Furthermore, changes in the link set require re-flooding the network link state even if the changed links were not being used. Since link state changes in wireless networks are often uncorrelated with data traffic and are instead caused by external (environmental) factors, this causes OSPF to fail both the control cost and loss response criteria. OSPF routers can impose policies on the use of links and can consider link properties (Type of Service), as the cost

associated with an edge is configurable by the system administrator [RFC2328], so receive a pass for link cost. However, there is no way to associate metrics with routers (as costs are only applied to outgoing interfaces, i.e. edges) when computing paths, and so fails the node cost criteria. While [RFC3630] discusses paths that take into account node attributes, it specifically states that no known algorithm or mechanism currently exists for incoporating this into the OSPF RFC.

IS-IS receives the same results as OSPF, because it maintains a consistent LSDB using similar mechanisms, and can account for link costs but not router costs in its shortest path computation.

"OLSRv2"

OLSRv2 is a proactive link state protocol, flooding link state information through a set of multipoint relays (MPRs). Routing state includes 1-hop neighbor information for each node in the network, 1-hop and 2-hop information for neighbors (for MPR selection), and a routing table (consisting of destination, and next hop), resulting in state proportional to network size and density (O(N*L + L^2)), and failing the routing state criterion.

Unacceptable control traffic overhead may arise from flooding and maintenance. HELLO messages are periodically broadcast local beacon messages, but TC messages spread topology information throughout the network (using MPRs). As such, control traffic is proportional to O(N^2). MPRs reduce this load to O(N^2 / L). As the number of MPRs is inversely proportional to the density of the network and L is bounded by N, this means control traffic is at best proportional to O(N).

Fisheye routing is a technique to reduce the frequency routing updates as the routing update propagates away from its source. This has the potential to reduce the control overhead to acceptable levels, and it is possible to implement this technique without violating the specification because the specification does not require that all updates be sent with the same frequency. However, there is no specification of how this should be accomplished. Thus, OLSR receives a "?" for the control traffic criterion. Fisheye routing does not alter the table size, so it does not modify OLSR's result ("fail") on the routing state criterion.

Furthermore, changes in the link set may require re-flooding the network link state even if those links were not being used by routing. OLSR makes these triggered floods optional, but as sending no triggered updates will raise problems in topology consistency, OLSRv2 receives a '?' in the loss response criterion.

OLSR allows for specification of link quality, and also provides a 'Willingness' metric to symbolize node cost, giving it a pass for both those criteria.

"TBRPF"

As a link state protocol where each node maintains a database of the entire network topology, TBRPF's routing table size scales with network size and density, leading to table sizes which are O(N * L) when a node receives disjoint link sets from its neighbors. This causes the protocol to fail the routing state criterion. The protocol's use of differential updates should allow both fast response time and incremental changes once the distributed database of links has been established. Differential updates are only used to reduce response time to changing network conditions, not to reduce the amount of topology information sent, since each node will periodically send their piece of the topology. As a result, TBRPF fails the control overhead criterion. However, its differential updates triggered by link failure do not immediately cause a global re-flooding of state (but only to affected routers) [RFC3684], leading to a pass for loss response.

TBRPF has a flexible neighbor management layer which enables it to incorporate various types of link metrics into its routing decision by enabling a USE_METRIC flag [RFC3684]. As a result, it receives a pass for link cost. It also provides a mechanism whereby routers can maintain multiple link metrics to a single neighbor, some of which can be advertised by the neighbor router [RFC3684]. Although the RFC does not specify a policy for using these values, developing one could allow TBRPF to satisfy this requirement, leading to a ? for the node cost requirement.

"RIP"

RIP is a distance vector protocol: all routers maintain a route to all other routers. Routing table size is therefore O(N). However, if destinations are known apriori, table size can be reduced to O(D), resulting in a pass for routing state. While standard RIP requires each node broadcast a beacon per period, and that updates must be propagated by affected nodes, triggered RIP only sends updates when network conditions change in response to the data path, so RIP passes the control cost criterion. Loss triggers updates, only propagating if part of a best route, but even if the route is not actively being used, resulting in a fail for loss response. The rate of triggered updates is throttled, and these are only differential updates, yet this still doesn't account for other control traffic (or tie it to data rate) or prevent the triggered updates from being flooded along non-active paths. [RFC2453]

RIP receives a ? for link cost because while current implementations focus on hop count and that is the metric used in [RFC2453], the RFC also mentions that more complex metrics such as differences in bandwidth and reliability could be used. However, the RFC also states that real-time metrics such as link-quality would create instability and the concept of node cost only appears as metrics assigned to external networks. While RIP has the concept of a network cost, it is insufficient to describe node properties and so RIP fails the node cost criterion..

"AODV"

AODV table size is a function of the number of communicating pairs in the network, scaling with O(D). This is acceptable and so AODV passes the routing state criterion. As an on-demand protocol, AODV does not generate any traffic until data is sent, and so control traffic is correlated with the data and so it receives a pass for control traffic. When a broken link is detected, AODV will use a precursor list maintained for each destination to inform downstream routers (with a RERR) of the topology change. However, the RERR message is forwarded by all nodes that have a route that uses the broken link, even if the route is not currently active, leading to a fail for loss response [RFC3561].

AODV fails the link cost criterion because the only metric used is hop count, and this is hardcoded in the route table entry, according to the RFC [RFC3561]. It fails the node cost requirement because there is no way for a router to indicate its [lack of] willingness to route while still adhering to the RFC.

"DYMO"

The design of DYMO shares much with AODV, with some changes to remove precursor lists and compact various messages. It still passes the routing state criterion because it only maintains routes requested by RREQ messages, resulting in O(D) table size. Control traffic (RREQ, RREP, and RREQ) are still driven by data, and hence DYMO passes the control cost criterion. The DYMO specification places very few requirements on how nodes respond to route error RERR messages that denote a broken route. Therefore, while it is possible for a DYMO implementation to meet the loss response criterion, the specification is not clear on how to meet the criterion while still maintaining routes as link breaks . This leads to a ? in loss repsonse [I-D.ietf-manet-dymo].

DYMO indicates that the "distance" of a link can vary from 1-65535 [I-D.ietf-manet-dymo], leading to a ? in link cost. While additional routing information can be added DYMO messages, there is no mention

of node properties, leading to a fail in node cost.

"DSR"

DSR performs on-demand route discovery, and source routing of packets. It maintains a source route for all destinations, and also a blacklist of all unidirectional neighbor links [RFC4728], leading to a total table size of O(D + L), failing the routing state criterion. Control traffic is completely data driven, and so DSR receives a pass for this criterion. Finally, a transmission failure only prompts an unreachable destination to be sent to the source of the message, passing the loss response criterion.

DSR fails the link cost criterion because its source routes are advertised only in terms of hops, such that all advertised links are considered equivalent. DSR also fails the node cost criterion because a node has no way of indicating its willingness to serve as a router and forward messages.

Appendix B. Logarithmic scaling of control cost

To satisfy the control cost criterion, a protocol's control traffic communication rate must be bounded by the data rate, plus a small constant. That is, if there is a data rate R, the control rate must O(R) + e, where e is a very small constant (epsilon). Furthermore, the control rate may grow logarithmically with the size of the local neighborhood L. Note that this is a bound: it represents the most traffic a protocol may send, and good protocols may send much less. So the control rate is bounded by O(R log(L)) + e.

The logarithmic factor comes from the fundamental limits of any protocol that maintains a communication rate. For example, consider e, the small constant rate of communication traffic allowed. Since this rate is communication, to maintain O(e), then only one in L nodes may transmit per time interval defined by e: that one node has a transmission, and all other nodes have a reception, which prevents them from transmitting. However, wireless networks are lossy. Suppose that the network has a 10% packet loss rate. Then if L=10, the expectation is that one of the nodes will drop the packet. Not hearing a transmission, it will think it can transmit. This will lead to 2 transmissions. If L=100, then one node will not hear the first two transmissions, and there will be 3. The number of transmissions, and the communication rate, will grow with O(log(L)).

This logarithmic bound can be prevented through explicit coordination (e.g., leader election), but such approaches assumes state and control traffic to elect leaders. As a logarithmic factor in terms

of density is not a large stumbling or major limitation, allowing the much greater protocol flexibility it enables is worth its small cost.

Authors' Addresses

Philip Levis
Stanford University
358 Gates Hall, Stanford University
Stanford, CA 94305-9030
USA

Email: pal@cs.stanford.edu

Arsalan Tavakoli
UC Berkeley
Soda Hall, UC Berkeley
Berkeley, CA 94707
USA

Email: arsalan@eecs.berkeley.edu

Stephen Dawson-Haggerty
UC Berkeley
Soda Hall, UC Berkeley
Berkeley, CA 94707
USA

Email: stevedh@cs.berkeley.edu