

On Performance Modeling of Ad Hoc Routing Protocols

Muhammad Saleem

*Center for Advanced Studies in Engineering
Islamabad 44000, Pakistan
Email: msaleem@case.edu.pk*

Syed Ali Khayam

*School of Electrical Engineering & Computer Science (SECS)
National University of Sciences & Technology (NUST)
Islamabad 44000, Pakistan
Email: ali.khayam@seecs.edu.pk*

Muddassar Farooq

*Next Generation Intelligent Networks Research Center (nexGIN RC),
National University of Computer and Emerging Sciences (FAST-NUCES),
Islamabad 44000, Pakistan.
Email: muddassar.farooq@nu.edu.pk*

Abstract

Simulation studies have been the predominant method of evaluating ad hoc routing algorithms. Despite their wide use and merits, simulations are generally time consuming. Furthermore, several prominent ad hoc simulations report inconsistent and unrepeatable results. We, therefore, argue that simulation-based evaluation of ad hoc routing protocols should be complemented with mathematical verification and comparison. In this paper, we propose a performance evaluation framework that can be used to model two key performance metrics of an ad hoc routing algorithm, namely routing overhead and route optimality. We also evaluate derivatives of the two metrics, namely total energy consumption and route discovery latency. Using the proposed framework, we evaluate the performance of four prominent ad hoc routing algorithms: DSDV, DSR, AODV-LL and Gossiping. We show that the modeled metrics not only allow unbiased performance comparison but also provide interesting insight about the impact of different parameters on the behavior of these protocols.

Key words:

Wireless ad hoc network, routing protocol, performance modeling, routing overhead, optimal route

1. Introduction

Since the inception of MANETs and sensor networks, a number of routing protocols have been proposed to efficiently discover and maintain paths in an ad hoc network [1, 2, 3, 4]. Due to lack of infrastructure and resources required to setup an ad hoc network, most of the routing protocols are evaluated and compared using network simulators. While simulations offer the flexibility to code and evaluate complex algorithmic logic, recent studies have shown that many ad hoc simulations report inconsistent, unrepeatable or incomplete simulation results [5]. Furthermore, scalability experiments

even with thousands of nodes quickly become infeasible due to extremely time consuming nature of network simulations. Therefore, we argue that simulation-based evaluation should be complemented with mathematical modeling of key performance parameters of ad hoc routing protocols. Such an approach will allow unbiased and provable performance comparison of the routing protocol even on very large scale networks. Moreover, an analytical approach will allow researchers to analyze the strengths and weaknesses of a protocol at an early design stage.

In this context, this paper proposes a performance evaluation framework to model the two most widely-

used performance metrics of wireless ad hoc routing algorithms [6], namely *routing overhead* and *route optimality*. We derive generic expressions to model the routing overhead incurred during the route discovery process. We also incorporate MAC layer channel contention and SNR-based physical and link layer channel error effects in the proposed modeling framework. For the route optimality model, we provide generic expressions to find the probability of optimal and suboptimal route discoveries for single as well as multipath routing protocols. We also model the expected probability of route establishment. Routing overhead and route optimality are the two baseline metrics and several other performance metrics may be derived from them. To illustrate the derivation process, we provide simple expressions for the following two metrics: *total energy consumption* and *route discovery latency*.

The derived metrics in the proposed framework are parameterized such that they can be adapted to specific ad hoc routing protocols. As a proof-of-concept, we use the proposed framework to model and compare four prominent ad hoc routing protocols, DSDV, DSR, AODV-LL¹ and Gossiping. To study the effect of modeling assumptions on the evaluated routing metrics, we also compare the analytical results with the results reported in an independently-conducted simulation study [6] and show that our modeling assumptions do not have a significant effect on the results. Finally, the proposed analytical models of routing protocols also yield interesting insight about the impact of different parameters on the performance of these protocols.

The rest of this paper is organized as follows. Section 2 summarizes the previous work in this area. Section 3 outlines system description, modeling assumptions and definitions of the key terms. The generic routing overhead model is developed in Section 4 along with the adaptation of model to specific protocols. Section 5 derives a model of route optimality in general and its application to specific ad hoc routing protocols. The use of the proposed models in mobile ad hoc networks is discussed in Section 6. The derivatives of routing overhead and route optimality are described in Section 7. We summarize key conclusions of this work in Section 8.

¹AODV-LL is a variant of AODV protocol [4] proposed in [6] that uses link layer feedback to avoid the use of HELLO messages to detect link failures.

2. Related Work

To counter the inherently unreliable nature of ad hoc networks, Tsirigos and Haas propose a routing scheme that makes use of multiple simultaneous paths [7]. With the addition of an overhead to each packet, the authors fragment the resulting unit into multiple blocks and route each one of them through a distinct path. The authors then study the probability of reconstructing the original information at the destination and show analytically that packet dropping probability decreases with an increase in the number of used paths. The authors then extend their work in [8] by relaxing the restrictions e.g. independent or disjoint paths, identical path failure probability. More specifically, the authors propose an approximation technique for selecting an optimal path set to maximize the probability of successful transmission in cases where paths may not be independent and their failure probability can vary as well.

Santivanez *et al.* [9] did scalability analysis of a number of ad hoc routing algorithms. The authors defined a scalability factor of routing protocol in terms of its total overhead and minimum traffic load. They showed that Plain Flooding (PF) algorithm scales better in high mobility scenarios while Hazy Sighted Link State (HSLs) [10] scales with the size of the network. Zhou and Abouzeid derive expressions for minimum length of control packets exchanged by cluster-based proactive ad hoc routing protocols [11]. The authors also provide scalability analysis of routing overhead with respect to network and cluster sizes. Expression for optimal cluster sizes to reduce the size of routing tables and the routing overhead are also derived by the authors. Xianren *et al.* propose a theoretical framework for computing the overhead of proactive routing protocols for MANETs [12]. The authors conclude that the overhead of the routing protocols rises with an increase in the node mobility.

Sadagopan *et al.* [13] developed a mathematical model to calculate the energy cost associated with a WSN query protocol called ACQUIRE. The authors also proposed similar models for Expanding Ring Search (ERS) and Flooding Based Queries (FBQ) for mutual comparison. Heusse *et al.* [14] theoretically analyzed the performance anomaly of 802.11b [15] by deriving expressions for throughput, probability of collision and contention time. The authors show that when some of the hosts communicate at lower bit rate, the performance of entire network degrades significantly. Jacquet and Laouiti [16] compares the performance of proactive routing protocols with flooding based routing mechanisms through simulations and the modeling of control overheads. However, it does not provide the fi-

nal analytical result and switches to simulation analysis. We proposed a performance evaluation framework for Bio-inspired ad hoc routing algorithms in [17]. However, the framework is only applicable to small-sized networks and does not cater for packet loss due to channel errors and collisions at the MAC layer.

One of the prominent analytical study is done by Bettstetter [18] in which a random graph model of ad hoc network topology is proposed. This model provides the node transmission radius, r_0 , that ensures a k -connected network for a given node density. We use Bettstetter's analysis as a foundation for the work reported in this paper.

3. System Description and Definitions

3.1. System description

We consider a dense ad hoc network of N nodes which are distributed on a two-dimensional plane using a homogeneous Poisson distribution with node density ρ . The resultant graph is assumed to be connected and all links (or edges) are symmetric. We assume an ad hoc network with a CSMA/CA-based MAC layer protocol for contention resolution. Even in the case of no contention, we account for the possibility that a packet may be lost due to channel errors (such as interference, fading etc.). For simplicity, we assume that the network topology does not change during a route discovery process. This assumption is realistic because timescale of mobility is much smaller than the timescale of a single route discovery. The routes may, however, change considerably from one route discovery to another. Similarly, we assume that the channel conditions do not change considerably during an RREQ transmission between two nodes. Finally, we assume that the nodes can transmit at a single uniform rate only.

3.1.1. Network topology

An ad hoc network is modeled as an undirected graph, $G = (V, E)$, where V is a set of nodes and E is a set of wireless links. A link between any two nodes exists if they lie within the transmission radius of each other. However, due to channel errors and contention, presence of a link between the neighboring nodes does not necessarily guarantee packets delivery in either direction.

A graph is connected if at least a single path exists between each pair of nodes; otherwise, it is a disconnected graph. A graph is said to be k -edge connected if and only if k edge-disjoint paths exist between each pair of nodes. $d(x)$ is the degree of a node x and it refers to

the number of nodes directly connected with x . A node is said to be an isolated node if it has a zero degree.

For the system described above, Bettstetter [18] derived an expression for the probability P that a randomly selected node has n_0 neighbors as given below.

$$P(d = n_0) = \frac{(\rho\pi r_0^2)^{n_0}}{n_0!} \cdot e^{-\rho\pi r_0^2}. \quad (1)$$

Putting $n_0 = 0$ in (1) gives the probability that a randomly selected node will be an isolated node; i.e., $P(d = 0) = e^{-\rho\pi r_0^2}$. Similarly, we can represent an average degree of a node as

$$E(d) = d_{avg} = \rho\pi r_0^2. \quad (2)$$

A network with $n \gg 1$ nodes is expected to be connected with probability p if r_0 is set to:

$$r_0 \geq \sqrt{\frac{-\ln(1 - p^{\frac{1}{n}})}{\rho\pi}}. \quad (3)$$

Equations (2) and (3) can be used to build the topology of an ad hoc network with given parameters. For instance, if we set $r_0 = 21$ m in a network of 10,000 nodes deployed in an area of 10^6 m^2 , the network will be connected with 99% probability.

3.2. Definitions

We now provide formal definitions of routing overhead and route optimality that are used in this paper.

3.2.1. Routing overhead and route optimality

Definition 1. Routing overhead of a reactive protocol is the total number of control packets launched in the network by the protocol in response to a route request (RREQ) message up to an arbitrary number of hops from the source node.

Optimal route is a protocol dependent term in which optimality may be defined in terms of the number of hops, energy, reliability, latency, etc. We use the most common and widely-used definition [2, 4].

Definition 2. An optimal route is a route of minimum hop length.

Definition 3. An n -suboptimal route between a \langle source, destination \rangle pair is a route of length $t + n$ hops, where t is the optimal length and $n = 1, 2, \dots$

Table 1: Description of Modeling Variables

<i>Symbols/category</i>	<i>Description</i>
Topology	
p_s	Broadcast forwarding probability
$p_r/\overline{p_c}$	Stochastic forwarding / no collision probability
$\overline{p_e}$	Probability of no channel error
d_{avg}/r_0	Average degree / Tr. radius of a node
N/ρ	Total number of nodes / node density
Routing overhead	
h	Number of hops from the source node
D	Diameter of the network (in hops)
T_s	Total simulation time
CW_{min}	Minimum contention window size in 802.11b
$SLOT$	Duration of a wait slot in 802.11b
t_{cont}	Contention time at a node
α	Rate of advertising routing table updates
$d_f[j]$	Exp. forward degree of a node at j hops
C_p	Generic routing overhead (packets)
$C_p^{(x)}$	Routing overhead of protocol x
Route optimality	
t	Length of an optimal path
k	Total number of node-disjoint paths
$w[i]$	Normalized weight of paths of length i
ϵ	Optimal path discovery probability
$X[t]$	No. of t hops paths discovered successfully
P	Probability of path discovery
$E\{X\}$	Expected path discovery probability (Generic)
μ	Number of tries to discover a link
Derivatives	
G_{total}	Total energy consumed
T_e/R_e	Tx / Rx energy per bit
$B_{rreq}/B_{rrep}/B_d$	Size of RREQ / RREP / data packets (bits)
M	Total number of data packets
L_a	Average path length
t_d	Route discovery latency
m	Number of paths discovered by a protocol

3.2.2. Broadcast forwarding probability

Nodes in ad hoc routing algorithms may forward RREQs stochastically. Additionally, the forwarded RREQs in these protocols may collide or get lost due to channel errors or dynamic topology. Therefore, we need not only to cater for the lost RREQs in the next hop calculations but also the stochastic forwarding behavior of nodes. For this purpose, we introduce a generic parameter, known as broadcast forwarding probability, defined as:

Definition 4. Broadcast forwarding probability is the probability that a node will forward a RREQ message to its neighbors and the message will be successfully delivered to them.

We can express the broadcast forwarding probability p_s as

$$p_s = p_r \cdot \overline{p_c} \cdot \overline{p_e}, \quad (4)$$

where $\overline{p_c}$ is the probability of *not* experiencing a collision at the MAC layer and $\overline{p_e}$ is the probability that the RREQ is *not* lost due to channel errors (We derive

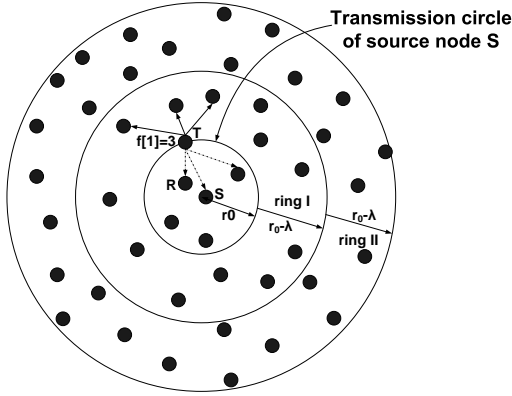


Figure 1: Illustration of the forward degree of a node.

expressions for p_c and p_e in Sections 4.1.2 and 4.1.3, respectively.) Finally, p_r is the probability with which a node will forward a RREQ to its neighbors; for instance, $p_r = 1$ for AODV-LL [4], while $p_r = p_g$ for the GOSSIP1(p_g) protocol [1]. (Further details on the stochastic forwarding probability p_r with reference to specific ad hoc routing protocols are provided later in the paper.)

3.2.3. Expected forward degree

Expected forward degree is relevant for understanding routing overhead model. Consider node T (Fig. 1) which has just received a RREQ and it decides to forward the RREQ to its neighbors. RREQ broadcast by node T will be received by its neighbors located in ring I and within the transmission circle of S . Now we must count those neighbors of T which are likely to forward this RREQ. In general, a neighbor drops a new RREQ if it already received the RREQ packet through other wireless links. Otherwise, it can decide to forward it with probability p_r . The neighbors of T in ring I that are expected to forward that RREQ define the forward degree of node T .

Definition 5. Expected forward degree of a node is the average (or mean) number of neighbors of that node which forward a received RREQ with probability p_s .

We represent the expected forward degree of nodes at j hops from the source node as $d_f[j]$ (see Table 1 for the description of modeling variables).

4. The Routing Overhead Model

In an attempt to quickly discover a route, reactive protocols completely or partially flood RREQs in a network. Tracking this flooding pattern is complicated because, depending on the routing scheme being used, the intermediate nodes may or may not forward the RREQs [1]. In this Section, we first model the routing overhead of reactive routing protocols as a function of expected forward degree which is later extended to proactive routing algorithms as well.

4.1. Generic routing overhead model in terms of expected forward degree

Consider source S that broadcasts RREQ to all its neighbors (Fig. 1). The first RREQ broadcast by S is received on average by $\overline{p_c} \times \overline{p_e} \times d_{avg}$ nodes. Each neighbor of S will forward the received RREQ with a probability p_r , therefore $p_s \times d_{avg}$ neighbors of S will forward RREQ to their neighbors. In this way, we keep accumulating the number of RREQs forwarded successfully at each hop up to an arbitrary number of hops (h) from the source node. Finally, the total expected routing overhead may be expressed as

$$C_p = \begin{cases} p_s d_{avg} & \text{if } h = 1 \\ p_s d_{avg} + d_{avg} \sum_{i=1}^{h-1} (p_s)^{i+1} \prod_{j=1}^i d_f[j]. & \text{otherwise} \end{cases} \quad (5)$$

See Appendix A for the derivation of (5). Equation (5) shows that C_p is directly related to the number of hops traversed by a RREQ, probability p_r and the expected forward degree of nodes. It can be intuitively validated as well. For instance, pure flooding algorithms have higher routing overhead because $p_r = 1$. Similarly, the routing overhead increases with an increase in path length between a given <source, destination> pair. Equation (5) provides an interesting insight that the RREQ storm exponentially decreases with an increase in the path length. As a result, a distant destination may not even receive RREQ and subsequently no path is discovered.

Now the routing overhead model (5) contains three unknown parameters; expected forward degree ($d_f[j]$), probability of no collision ($\overline{p_c}$) and the probability of no channel error loss ($\overline{p_e}$). The following three subsections model these three parameters.

4.1.1. Expected forward degree

Forward degree of a node depends upon its relative position from the broadcasting node. Generally speaking, the neighbors located at a longer distance from the broadcasting node have higher forward degree. For instance, node T (Fig. 1) can reach more new nodes than node R and therefore has a higher forward degree. Due to significant variance in the values of forward degrees, we derive expressions for an expected or mean forward degree of a node relative to its distance in terms of hops from the source.

Ideally, in the case of no channel error and contention, RREQ of node S (Fig. 1) is received by all the $\rho\pi r_0^2 - 1$ neighbors located within its transmission range. Broadcast of the neighbors of node S is received by nodes within $r_0 - \lambda$ radius of these neighbors where $0 \leq \lambda \leq r_0$. Note that λ is a function of node density and $\lambda \rightarrow 0$ for dense ad hoc networks. Therefore, for the presently-considered dense ad hoc network, the total number of *new* nodes that will hear the broadcast of the neighbors of S is approximately $3\rho\pi r_0^2$; i.e., nodes within ring I shown in Fig. 1. Note that the word *new* is used to factor out the other πr_0^2 nodes which have already received RREQ from the source node. This is necessary to incorporate the behavior of most ad hoc routing protocols in which nodes do not forward duplicate RREQs. Hence, the expected forward degree of 1 hop neighbors is $d_f[1] = \frac{3\rho\pi r_0^2}{\rho\pi r_0^2 - 1}$.

In the next iteration, nodes located in ring I will broadcast RREQ and cover ring II. Therefore, expected forward degree of 2 hop neighbors $d_f[2] = 5/3$. Broadcast storm continues to ripple across the network in the same manner. In general, expected forward degree of a node located at j hops from the source is $\frac{2j+1}{2j-1}$ where $1 < j < h - 1$. However, expected forward degree of nodes at $h - 1$ hops slightly differs from this formula. We combined all the three distinct cases in the following expression:

$$d_f[j] \approx \begin{cases} \frac{3\rho\pi r_0^2}{\rho\pi r_0^2 - 1} & \text{if } j = 1 \\ \frac{N - (h-1)^2 \cdot \rho\pi r_0^2}{(2h-3) \cdot \rho\pi r_0^2} & \text{if } j = h - 1 \\ \frac{2j+1}{2j-1}, & \text{otherwise} \end{cases} \quad (6)$$

where $1 \leq j \leq h - 1$. Remember that h in (6) represents the number of hops from the source node up to which we want to calculate the routing overhead. For protocols that flood the RREQs in the entire network e.g. AODV-LL, $h = \frac{D}{2}$ where D is the diameter of the region

measured in terms of hops. Equation (6) models a pure flooding protocol because we assume that broadcast of nodes within an inner region (e.g. ring I in Fig. 1) is heard in the entire outer region (ring II). In stochastic broadcasting protocols, however, we need to factor out those nodes that do not forward RREQ. With reduction in the number of forwarding nodes, potential receivers reduce proportionally in the outer ring. Therefore, we expect that the expected forward degree will still be the same because it is the ratio of nodes in the two rings.

The expected forward degree of nodes decreases with an increase in the number of hops (i.e. $d_f[1] > d_f[2] > \dots$). This is because of an increase in the number of potential forwarding nodes as the RREQ storm sways across a continuously decreasing uncovered area. We also emphasize that RREQ storm in real networks would not follow perfect ring styles as depicted in Fig. 1. Consequently, the routing overhead estimated by our model may be slightly higher than actual. A similar argument holds for sparse networks (or the cases where p_r is significantly smaller) where $\lambda > 0$.

4.1.2. Collision and contention modeling

Collision modeling. In our model, we assume standardized 802.11b [15] distributed coordination function (DCF) as a MAC layer protocol. Collisions happen in a distributed carrier sensing MAC scheme. Therefore, we must cancel out the propagation effect of collided RREQs. In a busy network where nodes always find the channel busy in the first attempt, probability of no collision \overline{p}_c is given by the following expression.

$$\overline{p}_c = \left(1 - \frac{1}{CW_{min}}\right)^{d_{avg}-1}, \quad (7)$$

where CW_{min} (= 31 as defined in 802.11b standard) is the minimum contention window and d_{avg} is the number of competing nodes. Higher is the average degree of a node, lower is the value of \overline{p}_c .

Contention modeling. Contention time is important because it contributes to the overall time consumed in a route discovery. We proceed with the same assumption as used in the derivation of \overline{p}_c . Since broadcast traffic is never retransmitted, contention time t_{cont} is

$$t_{cont} = \frac{SLOT \times \overline{p}_c \times CW_{min}}{2}, \quad (8)$$

where $SLOT$ (= $20\mu s$ for 802.11b standard) is the duration of each wait slot. For a given network, all variables in (8) are constants (d_{avg} , CW_{min} , $SLOT$) and hence a

node experiences a constant delay in order to get access to the channel for transmission of a RREQ.

4.1.3. Channel error modeling

We now briefly describe two prominent approaches that translate a channel model into the packet error probability p_e . These two approaches respectively involve physical and MAC layer channel models.

SNR-based physical layer channel error modeling. A commonly-used physical layer channel model for wireless ad hoc networks is the log-normal shadow fading model. This model characterizes the physical channel in a fading environment using two additive components: (1) a deterministic distance-dependent attenuation component with a path-loss exponent α , and (2) a SNR-based fading component defined as a normal random variable with a zero mean and variance σ^2 . The probability of a packet loss between two nodes communicating over this channel is given in [19]:

$$p_e = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{\beta_{th} - \alpha \times 10 \log(z)}{\sqrt{2}\sigma} \right), \quad (9)$$

where $\operatorname{erf}(\cdot)$ is the standard error function, z is the distance between the two nodes and β_{th} is the lowest threshold attenuation which is required to deliver a packet between the nodes. Probability of no channel error is then $\overline{p_e} = 1 - p_e$.

MAC layer channel error modeling. An alternative approach of modeling a wireless channel is to include physical layer errors as a part of the underlying channel model and then model the residual channel (i.e., the channel observed after physical layer processing) at the MAC layer. It has been shown that the bit-errors over a residual wireless channel exhibit bursty behavior where each successful and unsuccessful packet transmission is dependent upon whether or not the previous K packet transmissions were successful [20, 21]. Such a channel is adequately modeled using a K -th order Markov channel model in which the states of the model correspond to all possible combinations of correctly- and erroneously-received K previous packets. Then the total probability of packet error over a K -th order residual channel is:

$$p_e = \sum_{i=0}^{2^{K-1}-1} \pi_{2i+1}, \quad (10)$$

where π_i is the steady-state probability of Markov state i . Interested readers are referred to [20] and [21] for further details of Markov models of residual MAC layer

channels.

4.1.4. Discussion

This completes our routing overhead model. The final expression for routing overhead model is given below.

$$C_p = \begin{cases} p_r \overline{p_c} \overline{p_e} d_{avg} & \text{if } h = 1 \\ p_r \overline{p_c} \overline{p_e} d_{avg} + d_{avg} \sum_{i=1}^{h-1} (p_r \overline{p_c} \overline{p_e})^{i+1} \prod_{j=1}^i d_f[j]. & \text{otherwise} \end{cases} \quad (11)$$

In the following subsections, we discuss different optimizations to blind flooding of RREQs and the way they can be modeled using (11).

4.2. Optimized flooding of RREQs

Blind flooding of RREQs in the entire network may lead to broadcast storm problem [22]. Therefore, on-demand ad hoc routing protocols use a number of optimizations to this basic scheme. In the following subsections, we mention the most common optimizations that can easily be modeled using (11).

4.2.1. Ring search

This is one of the well-known methods in which source nodes set a *TTL* value in the broadcast RREQ indicating the number of hops it can travel. If $TTL > 0$, the receiving node(s) repeat the broadcast after decrementing the *TTL* value. Equation (11) provides a straight forward implementation of this process in which h represents the size of a ring. To illustrate further, we provide an expression for Dynamic MANET On-demand (*DYMO*) routing protocol [23] in Section 6 that uses an expanding ring search mechanism for the route discovery.

4.2.2. Stochastic flooding / gossiping

In this optimization, nodes forward RREQs with a certain probability p_r which may or may not vary with the hop distance. A number of variants of this approach are reported in [1]. In its most basic form, each node - irrespective of its distance from the source node - broadcasts RREQ with probability p_g . In other variations, the value of p_g varies for different nodes. For instance, the nodes within a radius of h_1 hops are allowed to broadcast unconditionally i.e. $p_g = 1$ while the rest do it with $p_g < 1$. Equation (11) can easily be adapted to any of these flooding patterns. To elaborate the transformation process, we adapt the generic routing overhead model to *GOSSIP1* protocol [1] in Section 4.3 that uses stochastic flooding of RREQs.

4.2.3. A hybrid approach

The optimizations techniques mentioned above may even be combined to further optimize the flooding process. For instance, ring search may be combined with stochastic flooding. It should now be evident that the proposed routing overhead model is able to handle such a hybrid case as transparently as the rest of the optimizations techniques.

In the following two subsections, we adapt the generic routing overhead model to four ad hoc routing algorithms AODV-LL, DSR, Gossiping and DSDV. Finally, we present a comparison of the results of routing overhead models of the four protocols with the ones obtained through simulations.

4.3. Protocol specific routing overheads

4.3.1. Dynamic Source Routing (DSR)

DSR [2] is an on-demand multipath ad hoc routing protocol with two main components; route discovery and route maintenance. During route discovery, source node broadcasts a RREQ packet to all its neighbors. Intermediate nodes keep rebroadcasting RREQs unconditionally (i.e. with $p_r = 1$) unless it reaches the destination which responds with a route reply (RREP). If an active route is available at an intermediate node, it also responds with a RREP. Route error (RERR) messages are used to inform the source node S that the current route is broken. AODV-LL [4, 6] route discovery mechanism is similar to that of DSR and hence routing overhead model of DSR also applies to AODV-LL.

As intermediate nodes in DSR forward RREQs with $p_r = 1$ in the entire network, (11) may be rewritten for DSR / AODV-LL as

$$C_p^{(dsr)} = \begin{cases} \overline{p_c p_e} d_{avg} & \text{if } h = 1 \\ \overline{p_c p_e} d_{avg} + d_{avg} \sum_{i=1}^{\frac{h}{2}-1} (\overline{p_c p_e})^{i+1} \prod_{j=1}^i d_f[j]. & \text{otherwise} \end{cases} \quad (12)$$

$\overline{p_c p_e}$ is the limiting factor in (12) that reduces the effective routing overhead. Note that the collided RREQs waste limited network resources like battery, bandwidth, etc., without contributing towards the route discovery process.

4.3.2. Destination-Sequenced Distance Vector (DSDV)

DSDV [3] is a proactive routing protocol in which each node maintains a next hop lying on the shortest path for all the reachable destinations in its routing table. In addition, it also stores its estimate of distance for

reaching the destination. A node periodically exchanges its distance vector routing tables with its neighbors. The loops are avoided by using sequence numbers in the exchanged messages.

Remember that all terms in (11) model the forwarding of RREQs which is not done in DSDV. Nodes only broadcast their routing table entries to their neighbors. Let the nodes exchange their distance vector table entries at a rate of α per second, total number of routing messages exchanged during an entire run of experiment is

$$C_p^{(dsdv)} = N \cdot T_s \cdot \alpha, \quad (13)$$

where T_s is the total time of an experiment. $C_p^{(dsdv)}$ has mostly constant parameters with the exception of α that can vary with changing topologies. As a result, the routing overhead of DSDV approximately remains constant with little variation under dynamic conditions [6].

Optimized Link State Routing Protocol version 2 (OLSRv2) [24] is another popular MANET routing protocol which performs restrictive flooding of Topology Control (TC) messages in a proactive manner. Speaking more specifically, OLSRv2 only allows specific nodes - called Multipoint Relays (MPRs) - to forward TC messages. Therefore, to model the routing overhead of OLSRv2, we have to combine (11) and (13). In this case, p_r will represent the probability that a given node is a MPR.

4.3.3. Gossip-Based Ad Hoc Routing

In Gossip based routing algorithms, each intermediate node rebroadcasts RREQ with a probability p_g , a technique termed as GOSSIP1(p_g) in [1]. The approach can be used to optimize a pure flooding protocol like AODV-LL. A number of variants of this basic approach are proposed in [1]. We only model a purely stochastic rebroadcasting approach in order to compare its routing overhead and route optimality characteristics with a plain flooding approach.

For GOSSIP1(p_g), (11) is modified to,

$$C_p^{(gossip)} = \begin{cases} p_g \overline{p_c p_e} \cdot d_{avg} & \text{if } h = 1 \\ p_g \overline{p_c p_e} \cdot d_{avg} \\ \quad + d_{avg} \sum_{i=1}^{\frac{h}{2}-1} (p_g \overline{p_c p_e})^{i+1} \prod_{j=1}^i d_f[j]. & \text{otherwise} \end{cases} \quad (14)$$

Equations (12) and (14) only differ by the values of probabilities. The product $p_g \overline{p_{ce}}$ in (14) is always less than $\overline{p_{ce}}$ irrespective of the individual values of p_g and $\overline{p_{ce}}$. As a result, $C_p^{(dsr)} > C_p^{(gossip)}$. But the lower value of the product $p_g \overline{p_{ce}}$ can significantly degrade the route optimality characteristics of GOSSIP1 as described in Section 5.

4.4. Comparison of routing overhead model with simulation results

We used a number of assumptions to simplify our routing overhead model, and therefore, we must compare the results of our model with the simulation results. In addition, we also evaluate the routing overhead model of DSR (equation (12)) in an ideal lossless channel $\overline{p_{ce}} = 1$. In this scenario, routing overhead of DSR generated by a source node should converge to the size of the network because no RREQs are lost. In a network of 200 nodes deployed in an area of 1000 m², each with a transmission radius of 3.97 m, our model estimated 199 RREQs that is in complete agreement with the expected results (see Table 2).

We used an impartial methodology for comparing the results of our model with the simulation results reported by the independent researchers (i.e. Broch *et al.* [6]). We generated the network topology as reported by the authors of [6] i.e. 50 nodes randomly placed in an area of 1500 m × 300 m. We used (3) to calculate the transmission radius of each node which equals 156.2 m. The simulation results [6] for 900 s pause time (in 30 sources scenario) and the corresponding formal model results are tabulated in Table 2.

One can see in Table 2 that the routing overhead estimated by our model is 4% less than that of DSR compared with NS-2 results. Keeping in view of the fact that we do not count the RERR messages nor the replies generated in response to a RREQ, 4% difference between the estimated routing overhead and the simulation results (AODV-LL and DSR) is acceptable. The reason for the matching results is due to slight over-estimation of the number of RREQs generated in a route discovery which indirectly balances the exclusion of RREPs, RERRs or HELLOs. It should also be noted that we are counting the routing overhead generated in a single route discovery attempt. This also contributes to the difference between the simulation and the analytical results.

We also repeated the same experiment for DSDV by assuming ($\alpha = 1, T_s = 900$) with results depicted in Table 2. Difference of 6.7% in the estimated result over 900 seconds is acceptable if we consider the dynamic

nature of wireless medium in ad hoc networks. Finally, we point out that the 0-order approximations (average values) used in the paper also lead to discrepancies between simulation and analytical results.

We also repeated the NS-2 experiments using the parameters reported in [6] to further elaborate the performance of routing overhead model. We disabled retries as our model only estimates the RREQs generated during a single route discovery. The results of these experiments are shown in the last row of Table 2. Now the results of the proposed model are extremely close to the simulation results. As expected, the model shows slightly higher routing overhead (see Section 4.1.1 for the reasons).

Before we move on to route optimality models, we would like to emphasize some other applications of the generic routing overhead model given by (11).

4.5. Other applications of the routing overhead model

While computing the routing overhead of ad hoc routing protocols, we in fact modeled a generic flooding pattern that is found in a variety of MANETs as well as WSN routing & data dissemination protocols. Our model estimates the number of nodes that may receive such a broadcast packet and are likely to repeat it after reception. Consequently, we are not only computing the total number of transmissions in the network but also compute the number of nodes that shall receive such a broadcast successfully. This in turn is a measure of the network coverage which is an extremely important parameter in ad hoc networks.

To elaborate this application, we modify (11) to compute the number of nodes that will receive the broadcast packet successfully N_s as given below.

$$N_s = \begin{cases} \overline{p_c p_e} d_{avg} & \text{if } h = 1 \\ \overline{p_c p_e} d_{avg} + p_c p_e d_{avg} \sum_{i=1}^{h-1} (\overline{p_r p_c p_e})^i \prod_{j=1}^i d_f[j]. & \text{otherwise} \end{cases} \quad (15)$$

Equation (15) shows that as a node distance (in hops) from the source gets higher, the probability that it will receive the packet reduces exponentially. Therefore, even if all nodes in the network take part in the broadcast process, it can not ensure that the broadcast will cover the entire network.

Table 2: Routing overhead: Simulation results reported by Broch *et al.* [6] Vs. the analytical results

No collision case		
Protocol	Analytical results	NS-2 results
AODV-LL/DSR	199	199
Simulation validation		
Protocol	Analytical results	Simulation results [6]
DSR / AODV-LL	1147	1200
DSDV	45000	42000
Protocol	Analytical results	Sim. results(0 retries)
DSR / AODV-LL	1147	1135

5. Route Optimality

5.1. Generic model

We now introduce our generic route optimality model for ad hoc routing protocols. In a typical ad hoc routing algorithm e.g. DSR, intermediate nodes do not forward duplicate RREQs and hence a node can be a part of single route only. Consequently, such ad hoc routing protocols only discover node-disjoint paths. Consider the network shown in Fig. 2 in which source node S floods a RREQ for destination D . Each intermediate node broadcasts a copy of RREQ until it reaches the destination. Now destination D may potentially receive as many RREQs as the number of its neighbors. In case of source routing protocols, RREQ builds a source routing header of the path followed by it. We have shown a copy of such RREQ received at D in Fig. 2. Now the question is: is it possible that a node may be present in the source routing headers of more than one RREQs received at D ? The answer is *no* because the presence of a node in multiple RREQs means that the node has rebroadcast multiple copies of identical RREQs which is not possible. Therefore, a node can only be a part of single route only supporting the argument that ad hoc routing protocols based on simple flooding can discover node-disjoint paths only.

We assume that there are k node-disjoint paths between a given <source, destination> pair with an optimal path length of t hops. We do not assume a fixed distribution of available paths. Rather, a function $f[i - t]$ provides the total number of node-disjoint paths of length i between a given <source, destination> pair. For instance, if there exist 6 node-disjoint optimal paths of length t and 3 node-disjoint paths of length $t + 1$ between the source and the destination, then $f[0] = 6$ and $f[1] = 3$.

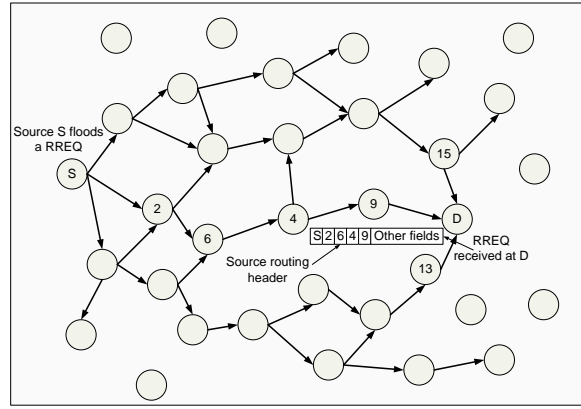


Figure 2: Route discovery in an ad hoc network: Source node broadcasts a RREQ which is flooded in the entire network.

5.1.1. Probability of optimal path discovery

The probability of discovering at least a single optimal path out of $f[0]$ optimal paths is

$$P(X[t] \geq 1) = 1 - (1 - \epsilon)^{f[0]\mu}, \quad (16)$$

where $\epsilon = (p_s)^t$ is the probability of discovering an optimal path and μ is the number of tries to discover a link. Derivation of (16) is given in Appendix B. The term $(1 - \epsilon)^{f[0]}$ in (16) is the probability that a routing algorithm fails to discover an optimal path in one attempt. We can minimize this probability either by increasing the number of optimal paths ($f[0]$) or by increasing p_s . This is easier said than done. For instance, increasing $f[0]$ by having a higher node density will increase the collision probability that results in decreasing the p_s . Note that the limiting case of $p_s = 1$ results in a zero failure probability.

5.1.2. Probability of suboptimal path discovery

We address this problem separately for single and multipath ad hoc routing algorithms.

Single path ad hoc routing protocols. In single path algorithms, discovering a 1-suboptimal route implies that an optimal path has *not* been discovered. A similar argument holds for 2-suboptimal route. We use this chain rule and (16) to come up with the following expression for the probability of discovering at least an n -suboptimal route.

$$P(X[t+n] \geq 1) = \left(1 - (1 - \epsilon(p_s)^n)^{f[n]\mu}\right) \times \prod_{j=0}^{n-1} \left(1 - \epsilon(p_s)^j\right)^{f[j]\mu}, \quad (17)$$

where $f[n-1]$, $f[n]$ provide the total number of available node-disjoint paths of length $t+n-1$ and $t+n$ hops respectively. In equation (17) as $n \rightarrow \infty$, the probability of finding n -suboptimal route approaches to zero. Therefore, suboptimal paths are less probable as compared to optimal paths.

Multiple path ad hoc routing protocols. Assuming that a multipath ad hoc routing protocol maintains all available paths, the probability of discovering an n -suboptimal path is

$$P(X[t+n] \geq 1) = \left(1 - (1 - \epsilon(p_s)^n)^{f[n]\mu}\right). \quad (18)$$

A simple comparison of (17) and (18) shows that the probability of discovering n -suboptimal paths in multipath routing protocols is greater than in single path routing protocols.

5.1.3. Expected probability of path establishment

Expected or marginal probability refers to the probability of discovering a path irrespective of the route length. It is important for two reasons: (1) ad hoc routing protocols are not always guaranteed to discover paths especially the ones that use stochastic rebroadcasting of RREQs and (2) expected probability of path establishment also provides us with an estimated path length that is discovered by a protocol. We again address this problem separately for the two classes of algorithms.

Single path ad hoc routing protocols. Using (16) and (17), the expected probability of path establishment

for a single path protocol is:

$$E\{X\} \simeq w[0] \left(1 - (1 - \epsilon)^{f[0]\mu}\right) + \sum_{i=1}^n w[i] \left(1 - (1 - \epsilon(p_s)^i)^{f[i]\mu}\right) \times \prod_{j=1}^i \left(1 - \epsilon(p_s)^{i-j}\right)^{f[i-j]\mu}, \quad (19)$$

where $w[i] = \frac{f[i]}{k}$ is the normalized weight of the paths of length $t+i$.

Multi-path ad hoc routing protocols. Using (16) and (18), the expected probability of finding a path in a multipath ad hoc routing protocol is

$$E\{X\} = \sum_{i=0}^n w[i] \left(1 - (1 - \epsilon(p_s)^i)^{f[i]\mu}\right). \quad (20)$$

Expected probability is a weighted average of all individual probabilities. As weights are dependent upon the number of available paths ($w[i] = \frac{f[i]}{k}$) of a particular length, expected probability degrades if the paths are evenly distributed. A higher value of expected probability can result in discovery of more optimal paths. A comparison of (19) and (20) shows that the multipath ad hoc routing algorithm has significantly higher probability of a route discovery as compared to single path algorithms (see Fig. 3(a) in which DSR has higher optimal path routing probability than AODV-LL).

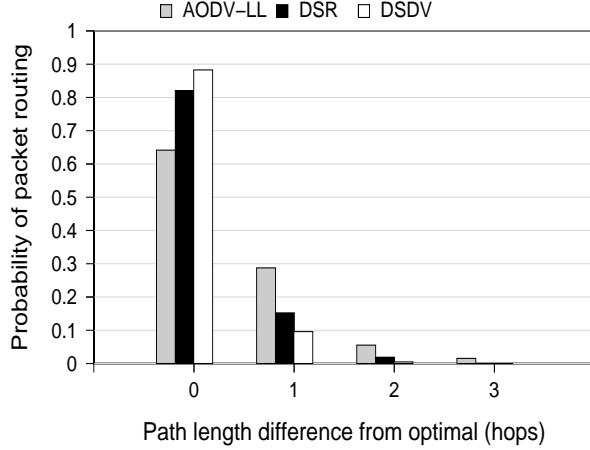
5.2. Protocol specific route optimality models

5.2.1. DSR and AODV-LL

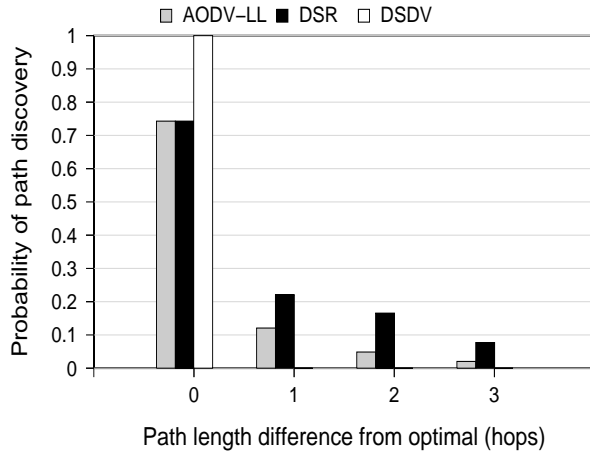
Intermediate nodes in both protocols forward the RREQs only once with $p_r = 1$ in a route discovery process. Therefore, putting $\mu = 1$ and $p_s = \overline{p_{ce}}$ in (19) and (20) gives expressions for marginal probability of path establishment for AODV-LL and DSR respectively. We do not reproduce these equations here for brevity. An interesting observation is that $\overline{p_{ce}}$ is a generic term that model both pure flooding and stochastic protocols in the same way. Therefore, both types of protocols shall fail to discover a destination located beyond a certain hop length.

5.2.2. DSDV

DSDV proactively maintains the shortest path to each possible destination. Each node in the network tries to discover a link with all its neighbors $T_s \cdot \alpha$ times in a given run of the experiment. Therefore, putting



(a) Packet routing probability through optimal / suboptimal paths [6]



(b) Optimal / suboptimal path discovery probabilities: analytical results.

Figure 3: A comparison of route optimality models with simulation results.

$\mu = T_s \cdot \alpha$ and $p_s = \overline{p_{ce}}$ in (19) gives the expected probability of route establishment for DSDV protocol. Note that the term $T_s \alpha$ increases the probability of optimal route discovery in DSDV. Consequently, the probability of discovering suboptimal paths is substantially reduced (see Fig 3(b)).

5.2.3. Gossip-Based Ad Hoc Routing

Gossiping is a strategy to control the flooding of RREQs. Therefore, it can be modeled both as single path and multipath routing protocol. In this model, we assume it to be a single path routing protocol in which the nodes forward only once the RREQs with probabil-

ity p_g in a route discovery process. Therefore, equation (19) can be modified for GOSSIP1(p_g) by inserting $p_s = p_g \overline{p_c} \overline{p_e}$ and $\mu = 1$. We assume in this model that even the source broadcasts a RREQ with probability p_g . The case in which source broadcasts a RREQ with probability 1 can be incorporated in (19) by putting $\epsilon = (p_s)^{t-1}$.

Since three probabilities in GOSSIP1(p_g) are multiplied i.e. $p_s = p_g \overline{p_c} \overline{p_e}$, therefore, the resulting probability is lower than the minimum value. If any one of p_g , $\overline{p_c}$ or $\overline{p_e}$ is small (e.g. $p_g < 0.7$), it will result in quick decaying of the broadcast storm. Consequently, a destination that is only fewer hops away from the source will not be discovered by the RREQs. It is therefore recommended that selection of rebroadcasting probability must be set carefully in order to make sure that RREQs are able to reach the farthest node in the network.

Before we conclude this section, we point out that although the adaptation process is elaborated for the expected route discovery probability models, generic models of optimal and suboptimal route discovery probabilities can also be adapted to specific protocols in a similar way.

5.3. Comparison of route optimality model with simulation results

The authors of [6] report the number of data packets routed by a protocol through optimal route, 1-suboptimal route, 2-suboptimal route and so on. To compare the results of the proposed route optimality models with simulation studies, we computed the probabilities that a given routing protocol routes a packet through optimal or suboptimal path. For this purpose, we have divided the number of data packets delivered through each route - optimal and suboptimal - by the total number of delivered packets. We plot the computed results in Fig. 3(a) for the three protocols: AODV-LL, DSR and DSDV. We used the same topology parameters as reported in [6] and the results of our framework for the three protocols are shown in Fig. 3(b). We assume $k = d_{avg}$, $t = 4$ and an exponential paths distribution i.e. $f[0] \gg f[1] \gg f[2] \dots$ and so on.

The bar graphs of simulation results and analytical results shown in Fig. 3 have approximately similar trends. Each protocol has higher optimal path probability while suboptimal paths are rarely discovered or used. The deviation from simulation results is expected due to our simplifying assumptions. A noticeable discrepancy in the analytical results is the longer tail of DSR. Recall that DSR route optimality model is based on the assumption of maintaining all available paths. Therefore,

the probability of suboptimal paths discovery is independent of the number optimal paths. However, implementation usually maintains a subset of the total available paths in which preference is given to the optimal paths. Consequently, suboptimal paths discovery probabilities decrease at a faster rate.

6. Evaluation of Ad Hoc Routing Protocols in Mobile Scenarios

The purpose of this section is two fold. First, we illustrate the use of the proposed modeling framework in mobile scenarios. Secondly, we elaborate the adaptation of our generic routing overhead model for *DYMO* protocol that uses an expanding ring search mechanism to control the RREQ flooding. The routing overhead model for *DYMO* is shown below.

$$C_p^{(dymo)} = \begin{cases} \overline{pcpe}d_{avg} & \text{if } \lceil \frac{d}{r_0} \rceil = 1 \\ \sum_{k=R_s}^{\lceil \frac{d}{r_0} \rceil} \overline{pcpe}d_{avg} + \sum_{k=R_s+1}^{\lceil \frac{d}{r_0} \rceil} d_{avg} \sum_{i=1}^{k-1} (\overline{pcpe})^{i+1} \prod_{j=1}^i d_f[j], & \text{if } \lceil \frac{d}{r_0} \rceil > 1 \text{ and } R_s = 1 \\ \sum_{k=R_s}^{\lceil \frac{d}{r_0} \rceil} \overline{pcpe}d_{avg} + \sum_{k=R_s}^{\lceil \frac{d}{r_0} \rceil} d_{avg} \sum_{i=1}^{k-1} (\overline{pcpe})^{i+1} \prod_{j=1}^i d_f[j], & \text{otherwise} \end{cases}$$

where d is the Euclidean distance between a given pair of <source, destination> nodes and R_s is the *initial ring size*. Keeping R_s as a variable parameter will allow us to study the impact of different values of R_s on the routing overhead.

Consider the network shown in Fig. 4 consisting of 50 nodes deployed randomly in an area of 500 m \times 500 m. The square around a node represents the area in which the node is allowed to move between two consecutive route discoveries. We simulate the movement of nodes by assuming a Poissonian RREQ generation pattern. Assuming squares of varying dimensions, we simulate different mobility patterns. We then performed experiments with different network topologies - 50, 200, and 1000 nodes - and different values of R_s . We report the total and the average routing overhead generated by a pair of <source, destination> nodes in Table 3.

It can easily be noticed that the number of RREQs generated under high mobility conditions are comparatively higher. Second important point to note is that the initial ring size R_s has significant impact on the routing

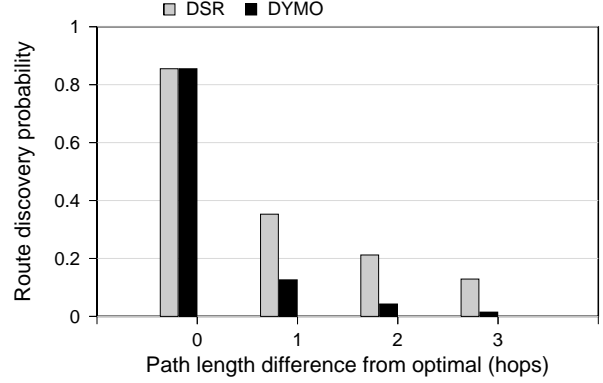


Figure 5: Route optimality of *DYMO* in mobile scenarios.

overhead. For instance, $R_s = 3$ produces the least routing overhead in all assumed scenarios. We also notice that too low values of R_s - in larger networks - can almost nullify the impact of ring search method. Consider a case in which a pair of <source, destination> nodes are separated by a distance of 5 hops. If the source node starts a route discovery process with $R_s = 1$, it will have to launch repeated route discoveries by incrementing R_s on every new attempt. Consequently, it will generate significantly higher routing overhead. On the other hand, if the source starts with $R_s = 3$, the number of attempts will be smaller. This clearly leads to an important conclusion that R_s should not be a constant value. Rather it should be set dynamically in accordance with node topology because the value of R_s is critically important to minimize routing overhead.

We have also computed the route optimality of *DYMO* and *DSR* protocol and the results are shown in Fig. 5. These results are consistent with the simulation results shown in Fig. 3 where the probability of optimal path discovery is significantly higher as compared to suboptimal paths. It then decays rapidly as the path length is increased.

7. Derivatives of Routing Overhead and Route Optimality

Route optimality and control overhead are the baseline metrics that can be used to derive a number of other metrics. As a proof of concept, in this section, we derive expressions for two additional metrics commonly reported for ad hoc routing protocols: energy consumption and route discovery latency. We use the symbols in Table 1 for the following derivations.

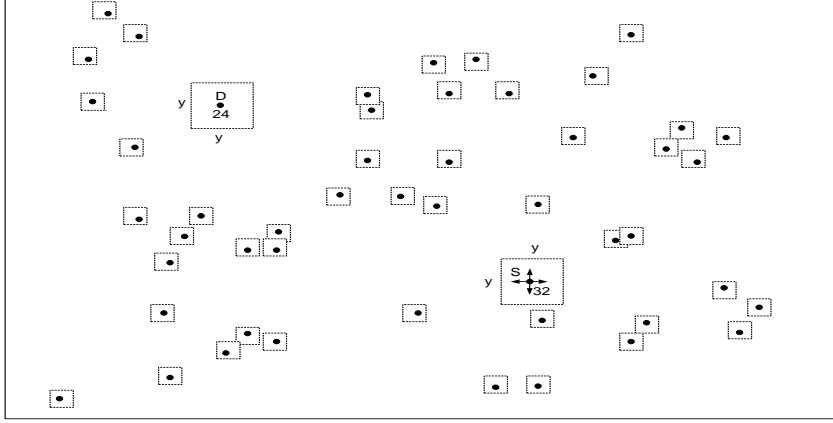


Figure 4: Example illustration of mobile scenarios.

Table 3: Routing overhead of *DYMO* under mobility

Number of nodes = 50, Area = 500m × 500 m	$R_s = 1$	$R_s = 2$	$R_s = 3$
Square size	Total (per RD)	Total (per RD)	Total (per RD)
100 m × 100 m	164 (41)	162 (41)	132 (33)
150 m × 150 m	201 (50)	199 (50)	168 (42)
200 m × 200 m	283 (55)	252 (49)	165 (33)
Number of nodes = 200, Area = 1000m × 500 m			
Square size	Total (per RD)	Total (per RD)	Total (per RD)
100 m × 100 m	346 (87)	344 (86)	334 (84)
150 m × 150 m	394 (99)	392 (98)	334 (84)
200 m × 200 m	540 (135)	512 (128)	405 (101)
Number of nodes = 1000, Area = 1000m × 1000 m			
Square size	Total (per RD)	Total (per RD)	Total (per RD)
200 m × 200 m	548 (137)	516 (129)	408 (102)
300 m × 300 m	619 (155)	587 (147)	479 (120)
400 m × 400 m	619 (155)	587 (147)	479 (120)

7.1. Energy consumption model

Transmission and reception of packets (control and data) are the major sources of battery depletion in mobile devices. We, therefore, derive expression for energy consumed during these two processes. Now the total energy consumed by a protocol is the sum of energy consumed during route discovery and data transmission phases.

If we know the expected route discovery probability of an ad hoc routing protocol (equation (19) & (20)), we can calculate the average length of the path (L_a) discovered by a protocol. Using the parameters listed in Table 1, final expression for the total energy consumed by a

protocol is

$$G_{total} = L_a(MB_d + mB_{rrep})(T_e + R_e) + C_p B_{rreq}(T_e + d_{avg}R_e). \quad (21)$$

An interesting point to note in (21) is that majority of energy consumed in a route discovery process is due to the reception of broadcast traffic which is approximately double the energy consumed for transmitting RREQs.

7.2. Route discovery latency

Route discovery latency is an important metric, especially for time critical applications. Using the average

path length discovered by a protocol (i.e. L_a) and the contention time given by (8), route discovery latency is given by

$$t_d = 2 \times L_a \times t_{cont}. \quad (22)$$

We ignore the transmission and propagation delays because of their dependence on technology but they can be embedded in (22). For a particular device type, the latency of route discovery is directly dependent on the path length. Higher the path length, more nodes contend for the channel access that lead to a higher latency for discovering routes.

8. Conclusions and Future Work

In this paper, we proposed a formal framework to accurately model key performance metrics of ad hoc routing algorithms. We adapted the generic models to specific ad hoc routing algorithms namely AODV-LL, DSR, GOSSIP1 and DSDV. We show that, in addition to time-efficient, accurate, consistent and repeatable performance evaluation of ad hoc protocols, the proposed modeling framework provides interesting insights into the impact of parameters on behavior of these protocols. Hence, we advocate the use of mathematical evaluation techniques for ad hoc routing algorithms because they are time-efficient, accurate, unbiased, consistent and scalable. In future, we shall extend the proposed evaluation framework in order to model other relevant metrics e.g. packet delivery ratio, packet latency etc. In addition to this, we also intend to make the framework more generic so that other emerging and prominent ad hoc routing protocols can be modeled and evaluated using different traffic patterns.

References

- [1] Z. J. Haas, J. Y. Halpern, and L. Li. Gossip-based ad hoc routing. *IEEE/ACM Transactions on Networking*, 14(3):479 – 491, June 2006.
- [2] D. B. Johnson, D. A. Maltz, and J. Broch. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*, edited by Tomasz Imielinski and Hank Korth, page 153181, 1996.
- [3] C. E. Perkins and P. Bhagwat. Highly dynamic destination-sequenced distance-vector routing (dsv) for mobile computers. In *Proc. of SIGCOMM 94*, Aug., 1994.
- [4] C. E. Perkins and E. M. Royer. Ad-hoc on-demand distance vector routing. In *Proc. of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, Feb., 1999.
- [5] S. Kurkowski, T. Camp, and M. Colagrosso. MANET simulation studies: The incredibles. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9(4):50 – 61, October 2005.
- [6] J. Broch, D. A. Maltz, D. B. Johnson, Y. C. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proc. of ACM/IEEE MOBICOM'98*, volume 1, pages 85–97, Dallas, Texas, USA, October 1998.
- [7] A. Tsirigos and Z. J. Haas. Analysis of multipath routing-part i: the effect on the packet delivery ratio. *IEEE Transactions on Wireless Communications*, 3(1):138 – 146, Jan 2004.
- [8] A. Tsirigos and Z. J. Haas. Analysis of multipath routing, part 2: Mitigation of the effects of frequently changing network topologies. *IEEE Transactions on Wireless Communications*, 3(2):500 – 511, March 2004.
- [9] C. A. Santivanez, B. McDonald, I. Stavrakakis, and R. Ramanathan. On the scalability of ad hoc routing protocols. In *Proc. of IEEE INFOCOM*, 2002.
- [10] C. Santivanez, S. Ramanathan, and I. Stavrakakis. Making link state routing scale for ad hoc networks. In *Proc. of MobiHoc*, Long Beach, CA, Oct. 2001.
- [11] N. Zhou and A. A. Abouzeid. Routing in ad hoc networks: A theoretical framework with practical implications. In *Proc. of IEEE INFOCOM*, pages 1240– 1251, Miami, Florida, USA, March 2005.
- [12] W. Xianren, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves. Routing overhead as a function of node mobility: Modeling framework and implications on proactive routing. In *Proc. of IEEE International Conference Mobile Adhoc and Sensor Systems*, pages 1–9, Montreal, Canada, Oct. 2007.
- [13] N. Sadagopan, B. Krishnamachari, and A. Helmy. Active query forwarding in sensor networks. *Elsevier Ad Hoc Networks*, 3:91–113, January 2003.
- [14] M. Heusse, F. Rousseau, G. B. Sabbatel, and A. Duda. Performance anomaly of 802.11b. In *Proc. of IEEE INFOCOM*, pages 836– 843, San Francisco, USA, April 2003.
- [15] Wireless LAN medium access control (MAC) and physical layer (PHY) specification, 2000.
- [16] P. Jacquet and A. Laouiti. Overhead in mobile ad-hoc network protocols. INRIA Research report 3965, 2000.
- [17] M. Saleem, S. A. Khayam, and M. Farooq. A formal performance modeling framework for bio-inspired ad hoc routing protocols. In *ACM GECCO '08*, July, 2008.
- [18] C. Bettstetter. On the minimum node degree and connectivity of a wireless multihop network. In *Proc. of MobiHoc*, pages 80 – 91, Lausanne, Switzerland, June 2002.
- [19] C. Bettstetter and C. Hartmann. Connectivity of wireless multihop networks in a shadow fading. *ACM/Kluwer Wireless Networks Journal*, 2005.
- [20] S. A. Khayam and H. Radha. Linear-complexity models for wireless mac-to-mac channels. *ACM/Kluwer Wireless Networks Journal*, 2005.
- [21] A. Konrad, B. Y. Zhao, A. D. Joseph, and R. Ludwig. A markov-based channel model algorithm for wireless networks. *ACM/Kluwer Wireless Networks Journal*, 2003.
- [22] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proc. of ACM/IEEE MobiCom*, pages 151 – 162, Washington, USA, 1999.
- [23] I. Chakeres, and C. Perkins. Dynamic MANET On-demand (DYMO) Routing. *Internet-Draft*, <http://ianchak.com/dymo/draft-ietf-manet-dymo-12.txt>, 2008.
- [24] T. Clausen, C. Dearlove, and P. Jacquet. The Optimized Link State Routing Protocol version 2. *Internet-Draft*, <http://potaroo.net/ietf/all-ids/draft-ietf-manet-olsrv2-10.txt>, 2009.

A. Derivation of Generic Routing Overhead Expression

Equation (5) uses the concept of expected forward degree ($d_f[j]$). Multiplying $d_f[j]$ with the total number of rebroadcasting nodes at j hops yields the number of nodes that are likely to forward RREQ at $j+1$ hops. For example, multiplying $p_s \cdot d_{avg}$ (rebroadcasting nodes at 1 hop) with $d_f[1]$ gives us the maximum number of nodes at two hops from the source that may forward the RREQ broadcast ($p_s d_{avg} d_f[1]$). Including p_s in $p_s d_{avg} d_f[1]$ factors out lost RREQs as well as the nodes that do not participate in broadcast forwarding. Therefore, RREQs generated at two hops is $p_s(p_s d_{avg}) \times d_f[1]$, the third term in (23). We sum up all the RREQs generated at each hop and get

$$\begin{aligned}
 C_p = & 1 + p_s \cdot d_{avg} + p_s (p_s \cdot d_{avg}) \times d_f[1] + \\
 & p_s (p_s^2 \cdot d_{avg} \cdot d_f[1]) \times d_f[2] + \\
 & p_s (p_s^3 \cdot d_{avg} \cdot d_f[1] \cdot d_f[2]) \times d_f[3] + \dots + \\
 & p_s (p_s^{h-1} \cdot d_{avg} \cdot d_f[1] \dots d_f[h-2]) \times d_f[h-1] \\
 & -1,
 \end{aligned} \tag{23}$$

where $d_f[1], d_f[2], d_f[3], \dots, d_f[h-1]$ represent the expected forward degree of nodes at 1, 2, 3, ..., $h-1$ hops from the source node respectively. Also notice the last 1 that is subtracted from the sum of all terms. This indicates that the destination node does not participate in rebroadcast. Equation (5) is the concise form of (23).

B. Derivation of Probability of Optimal Path Discovery

Remember that the optimal paths are t hops long and each link is discovered with a probability p_s , therefore, probability of discovering an optimal path is $\epsilon = (p_s)^t$. Hence, the probability of failing to find an optimal path is $(1 - \epsilon)$. The current problem represents a Bernoulli trial with ϵ as the probability of success and $(1 - \epsilon)$ as the probability of failure. Therefore, the probability of finding j optimal paths out of a total of $f[0]$ optimal paths is a binomial distribution given by

$$b(j; f[0], \epsilon) = P(X[t] = j) = \binom{f[0]}{j} \epsilon^j (1 - \epsilon)^{f[0]-j}, \tag{24}$$

where $X[t]$ is a random variable representing the number of optimal paths discovered successfully. Putting $j = 0$ in (24) gives the probability that no optimal paths will

be discovered. Hence, the probability of discovering at least one optimal path is $1 - P(X[t] = 0)$ as given in (16).