Ann. Telecommun. (2012) 67:269–284
DOI 10.1007/s12243-011-0272-9

A routing protocol using a reliable and high-throughput path metric for multi-hop multi-rate ad hoc networks

Hieu Cao Trong · Sungwon Lee · Choong Seon Hong

Received: 9 March 2010 / Accepted: 8 July 2011 / Published online: 27 August 2011

Abstract In this paper, a high-throughput routing protocol for multi-rate ad hoc networks using lower layer information is proposed. By choosing the route with the minimum value of the proposed "Route Assessment Index" metric which has the form of entropy function, the selected route is ensured to have high throughput and link reliability among route candidates. Link bottleneck is avoided in the chosen route; hence, the packet drop rate due to buffer overflow is alleviated. Furthermore, an effective route discovery strategy is also introduced along with new routing metric. The correctness of the proposal is proven, and the simulation results show that our new metric provides an accurate and efficient method for evaluating and selecting the best route in multi-rate ad hoc networks.

Keywords Multi-rate ad hoc networks · High throughput · Reliability · Routing metric · Cross layer

1 Introduction

Ad hoc networks currently have become an ideal topology for establishing instant communication infrastructure where other kinds of networks have difficulties to be deployed. Each node has the ability to communicate directly with any other in its communication range, while the out-of-range nodes use intermediary hops to communicate with each other. The wireless ad hoc networks (including wireless sensor networks) are applicable to a wide variety of fields as they are operable without any predefined infrastructure.

Nowadays, physical layer enhancements support multiple data rates, which enables wireless nodes to select the appropriate transmission rate depending on the required quality of service and the radio channel conditions. For example, the IEEE 802.11g standard [13] with orthogonal frequency division multiplexing (OFDM) technology support eight modulation and coding schemes (MCS) and offers eight data rates between 6 to 54 Mbps according to the selected MCS, as shown in Table 1.

Considering the multi-rate ad hoc networks, although there are a huge number of papers proposed a numerous routing solutions for wireless networks with different approaches, up to now, there are still very few papers considering design an effective routing metric that can utilize the benefit of using multi-rate. The main reason is most of the papers gave solutions only in routing layer; hence, the helpful parameters from MAC layer were not used to design a more effective routing protocol. Regarding this point, there are some papers using lower layer information for routing decisions. However, only a few well-known existing literatures have a comprehensive routing metric and protocol design, while others are the derivative solutions that only solved a small aspect of routing problems. The selected comprehensive contributions are weighted cumulative expected transmission time (WCETT) [9] and medium time metric (MTM) in [3]. As discussed in [3], there is a direct relationship between the rate of communication

H. Cao Trong · S. Lee · C. S. Hong (✉)
Department of Computer Engineering,
Kyung Hee University, Seoul, South Korea
e-mail: cshong@khu.ac.kr

H. Cao Trong
e-mail: hieuct@khu.ac.kr

S. Lee
e-mail: drsungwon@khu.ac.kr

Table 1 Data rate and Rx sensitivity in 802.11 OFDM PHY

Data rate r_k (Mbps)	Modulation type	Coding	Rx sensitivity $P_{S_{r_k}}$ (dBm)
06	BPSK	1/2	−82
09	BPSK	3/4	−81
12	QPSK	1/2	−79
18	QPSK	3/4	−77
24	16-QAM	1/2	−74
36	16-QAM	3/4	−70
48	64-QAM	1/2	−66
54	64-QAM	3/4	−65

and the transmission range. Since distance is one of the primary factors that determines wireless channel quality, there is an inherent trade-off between high transmission rate and effective transmission range. A low rate link can cover the distance to the destination in few hops, while a high rate link requires more hops to reach the destination. As such, high rate route must deal with more risk of broken links and route discovery delay, due to the extra hops to the destination.

In this paper, we give a comprehensive proposal for routing metric and protocol design in the multi-rate ad hoc networks. Considering both the end-to-end throughput of a route and the rate of each link corresponding to relative distance of two nodes to avoid selecting bottleneck links, we proposed a new routing metric named "Route Assessment Index" (RAI) for multi-rate ad hoc networks. The proposed metric has the form of entropy function. Therefore, the properties of entropy function are used to show the expected characteristics of a communication route, such as small number of hop, no link bottleneck, and high throughput. Also, since wireless links are not completely reliable, the routing metric uses the link reliability information (packet delivery rate) from the MAC layer to choose the best route. The route with minimum RAI value is proved to be the high-throughput and link reliability route. That route also has a small number of intermediate nodes among route candidates. Therefore, the end-to-end throughput increases significantly. In this paper, we use the term *link reliability* to refer the ability of a link to successful deliver data packets and the term *effective link capacity* to refer the combination of link reliability and link rate. The detail will be showed in Section 3.2.

The remainder of this paper is organized as follows: In Section 2, we analyze related work and some well-known routing metrics. The main part including proposed protocol's model and operation with an example is presented in Section 3. The performance of RAI is given in Section 4. Finally, in Section 5, we conclude our paper.

2 Related work

A lot of routing protocols have been proposed for the (mobile) wireless ad hoc networks, which follow one of two major strategies: proactive one such as in DSDV [17] and OLSR [5], or a reactive (on-demand) one, such as in ad hoc on demand distance vector (AODV) [16] and DSR [14]. These protocols were originally designed for single-rate networks, thus have used a shortest path algorithm with minimum hop count metric to select paths. Min hop is a good metric in single rate networks where all links are equivalent. However, it does not perform well in the multi-rate wireless network because it does not utilize the higher link rate for data transmission.

The AODV protocol [16] is one of the popular reactive routing protocol that discovers the route between the source and destination nodes dynamically. In AODV, when the source node wants to communicate with a destination node, it will broadcast a route request (RREQ) packet to the network. The neighboring nodes, which receive the RREQ packet, search for an existing route to the destination in its routing table. If a route already exists, then the intermediate node replies with an unicast route reply (RREP) packet to the RREQ sender. Otherwise, the node forwards the RREQ packet to its neighbors. By this way, the RREQ packet traverses hop by hop and reaches the destination. The destination node replies with an RREP to establish a new route by sending the packet traversing the same path in the reverse direction. When the source node receives multiple copies of RREP packets for the same RREQ packet, it selects the route with the minimum number of hops. The hello and route error (RERR) packets are used to manage route failure and reconstruction. The design of the AODV protocol is based on the simple packet radio model without the consideration of data transmission rate. The main problem of AODV is based on hop count, which may avoid to choose the highest data rate route.

The automatic rate fallback (ARF) protocol originally developed in [15] has been widely adopted by the industries to determine the initial transmission rate. In ARF, the node first transmits packet to a particular destination at the highest data rate, and it switches to the next available lower data rate when it does not receive two consecutive ACK frames and starts a timer after the switch. When the node receives ten consecutive ACK frames successfully or the timer expires, it switches to the next higher data rate again and packets are always transmitted at the highest possible rate. In another paper, the receiver-based auto rate (RBAR) protocol [11] allows the receiving node to select the rate. This is accomplished by using the signal-to-noise ratio of the request-to-send packet to choose the most appropriate rate. The clear-to-send packet is used to ACK that rate to the sender. The opportunistic auto rate (OAR) protocol presented in [19] operates using the same receiver based approach. It allows high-rate multi-packet bursts to take advantage of the coherence times of good channel conditions. OAR uses the IEEE 802.11 mandated fragmentation field to hold the channel for an extended number of packet transmissions. In IEEE 802.11, each node has equal opportunity to send the same number of packets; hence, the node transmitting at a high rate actually does not gain high throughput if it shares the channel with some nodes at lower transmission rate. However, in OAR, each node accesses the medium for the same amount of time, so the overall throughput will increase with the higher link rates. Therefore, both RBAR and OAR require modifications to the 802.11 standard but can increase the overall throughput.

For multi-rate wireless ad hoc networks, several routing metrics were proposed. The author in [10] introduced an approach for multi-rate MANETs to improve traditional AODV routing protocol. The proposal was based on the link cost which is simply provided by delay time for transferring a packet from MAC layer which is inherited from [2]. The paper is a derivative one of MTM paper. Therefore, in this section, we present and analyze two well-known routing metrics which were comprehensive designed for multi-rate wireless ad hoc networks.

2.1 Weighted cumulative expected transmission time

Draves et al. [9] proposed multi-radio link-quality source routing (MR-LQSR) protocol for multihop wireless network that uses multi-rate. MR-LQSR is a combination of the LQSR protocol [8] and a routing metric WCETT. The motivation of this work is the consideration of link bandwidth and PHY-layer loss rate, which both affect packet transmission time. Also, it prefers channel diversity.

First, the expected transmission time (ETT) metric was proposed in Eq. 1. They calculated the time required to transmit a packet of size S on a link with a data rate B (raw data rate) under the number of transmission (including retransmissions) attempts ETX [6]. The ETT of a link is the duration of time a node uses the medium to successfully deliver a packet to the next hop. Thus, ETT of the i-th link is defined by

$$\mathrm{ETT}_i = \mathrm{ETX}_i \times \frac{S}{B_i}, \tag{1}$$

where B_i is the data rate of the i-th link. Here, ETX is calculated as

$$\mathrm{ETX} = \frac{1}{d_f \times d_r}, \tag{2}$$

where d_f and d_r denote the delivery ratio in the forward and reverse directions, respectively. Next, to find routes with less intra-flow interference and channel diversity, the authors in [9] proposed a new routing metric, which is defined by

$$\mathrm{WCETT} = (1-\beta) \times \sum_{i=1}^{n} \mathrm{ETT}_i + \beta \max_{1\le j\le k} X_j, \tag{3}$$

where X_j is the summation of ETT of the links in a path p operating on channel j, k is the number of orthogonal channels available, and $0 \le \beta \le 1$ is a tunable parameter. Equation 3 can be interpreted as a balance between the end-to-end delay experienced in a particular route (the first component) and the channel diversity (the second component). MR-LQSR was designed to use multiple radios. For the testbed, each node has two different wireless cards. One card operated in IEEE 802.11a on channel 36, and the other operated in IEEE 802.11g on channel 10 statically. Both cards used autorate. However, MR-LQSR cannot avoid link bottleneck and may choose route with more hops.

2.2 Medium time metric

Awerbuch et al. [3] showed the efficiency of the MTM in selecting high-throughput route. MTM uses the total medium time of a packet in a given path, where the medium time is defined as the time needed to transmit a packet on a given link with a particular data rate including the MAC delays and control overheads.

The simulation results in [3] showed the relationship between the throughput across the path and the length of the path:

1. At certain distances, low rate links can achieve higher throughput than high rate links because high rate path may take more hops.
2. Due to spatial reuse, as the path becomes longer, multiple transmissions can take place along the path at the same time.
3. High rate links can achieve high throughput after this distance though more hops needed.

The authors claim that MTM can select optimal throughput paths and tends to avoid long unreliable links. MTM assigns a weight to each link in the path, which is proportional to the packet transmission time on that link, and then adds all the weights for the path.

When applying MTM to on-demand routing protocols such as DSR, it will result in the path lasting longer. The proactive routing protocol DSDV [17] is modified by using MTM as metric instead of hop count. It also uses OAR as the lower layer to provide multi-rate access and the current communication rate. The strong point of MTM is simplicity. It only needs the link rates provided by OAR instead of link utilization which is difficult to detect. The simulation results show that by combining MTM and OAR, throughput gains of up to 100% to 200% can be achieved over traditional route selection. However, the weak points of MTM are it was designed for proactive routing protocol so that each node must maintain the whole network information of others. Without an effective route discovery strategy, the needed time to find a good route is considerable. Also, MTM was not designed to avoid selecting some particular lower rate links; hence, they will cause high packet drop rate at those bottleneck links. Consequently, the throughput will be downgraded. Our proposed routing protocol will solve those weaknesses.

Recently, there are several most up-to-date papers deal with cross-layer approach for the effective routing with some interesting techniques [4, 12, 20, 22]. Kai Hong et al. [12] proposed a cross-layer MAC design, which improves the coordination between MAC and routing layers using an idea called "virtual link" for the wireless ad hoc networks. They claimed that none of work focuses on the long processing delays in a relay/forwarding node due to the layer structure in wireless communication; hence, the energy consumption for every data frame is also increased due to long processing delays. To solve that problem, the "virtual link" is used, and the concept of "virtual link" can be defined as follows: Suppose node B is an intermediate node of node A and C (A and C are out of communication range, alternatively, they do not have physical link), but note that physical links exist between A–B and B–C. If these two physical links could be combined to obtain a "virtual link", then we obtain a logical link between A and C. However, to obtain the list of virtual links, the proposed MAC protocol must establish and maintain virtual links automatically and periodically. In addition, in the paper, the concept of virtual link works with "inbound monitor" and "self-learning" modules, which also need to pre-process some information such as monitor physical links, virtual links, and even re-encapsulate the packets. Those processes actually are not simpler and even more complicated than the traditional steps of existing routing protocol (checking IP address, choosing the next hop based on desired parameters, forwarding packets to the next hop, etc.). Especially, the paper did not show any specific metric or parameter which can be used to assess the created virtual links and to select the best virtual link among possible candidates. Considering the interference in the wireless communication as the main factor, authors in [4] give an interference-aware multi-path routing and link rate control for wireless networks. The paper used a new approach with cross-layer interactions, in which flows can split or merge at any node, and there is no preselected route or routing policy other than to maximize throughput. A routing path is constructed based on the link rate allocated for each flow; hence, the authors have used a different approach with traditional routing, which selects a route for the two end nodes before sending data. For wireless multimedia sensor networks, a cross-layer design was proposed in [20]. The desired target is to maximize the number of video stream requests to be delivered while the imposed distortion constraint on the streams are met. The paper only focused on the QoS issues for a specific network. For the multichannel wireless mesh networks, [22] address the problem of joint routing and transmission scheduling with variable-width channel allocation. While narrower bands split the total available spectrum into more non-overlapping channels allowing more parallel concurrent transmissions, wider bands increase the capacity of the communication links. The paper has shown that variable-width channel assignment achieves significant improvement over fixed-width allocation. Traffic of a particular session may be split to sub-flows routed over different paths and assume a time division multiple access scheme where time is divided into slots and a link may be active in one or more time slots to meet the traffic requirement. Several links may be active in the same time slot without violating the signal-to-interference and noise ratio requirement. This solution has the same approach with our proposal in the previous work [21], where multi-path routing and scheduling are jointly cooperated. However, all above existing work in cross-layer design uses different approaches with conventional metric-based routing which is used in our paper. Therefore, we only focus on the comparisons of our work with metric-based routing protocols in this paper.

3 Proposed routing protocol

In this section, we will discuss about the relation between transmission range and communication rate based on the received *received signal strength indicator (RSSI)*. Then, we propose the new routing metric working for reactive protocol. Along with routing metric, the route discovery phase also plays an important role in a routing protocol by helping source find as good as possible route to destination within negligible time. Therefore, in this section, we introduce a new strategy for route discovery. Finally, we explain the operations of our routing protocol and give an example to clearly demonstrate the operations of our protocol.

3.1 Relation between transmission range and communication rate

As modeled and calculated in many existing literatures, the channel quality is the main factor to decide which transmission rate can be used. For transmitting data at a specific rate r_k (i.e., r_k = (6, 9, 12, 18, 24, 36, 48, 54) Mbps), the corresponding receiver sensitivity is required as showed in Table 1. Remember that the number of rate levels as well as the maximum data rate here follow the IEEE 802.11g standard [13]. Table 1 shows the data rate and Rx sensitivity in IEEE 802.11 OFDM PHY.

Hence, to transmit data at rate r_k, the received signal strength must at least equal the receiver sensitivity $P_{S_{r_k}}$. Using the log-distance path loss model in [18] for radio propagation, the received signal strength at receiver R with distance d far away from the transmitter T is calculated as:

$$P_r = P_t - 20\log_{10}\left(\frac{4\pi \overline{d} f}{c}\right) - 10\gamma \log_{10}\left(\frac{R_{r_k}}{\overline{d}}\right) \text{ (dBm)} \tag{4}$$

in which P_r and P_t are the receive and transmit signal power in dBm, $20\log_{10}\left(\frac{4\pi \overline{d} f}{c}\right)$ is the free space path loss at a reference distance $\overline{d}$ (normally, 1 m) in dBm for signal speed of c and frequency f, and γ is the path loss exponent ($2 \leq \gamma \leq 6$) depending on the environment condition between T and R. From Eq. 4, let $P_r = P_{S_{r_k}}$ and $\overline{d} = 1$, we can determine the transmission range R_{r_k} at rate r_k as:

$$R_{r_k} = 10^{\frac{P_t - P_{S_{r_k}} - 20\log_{10}(4\pi f/c)}{10\gamma}} \tag{5}$$

As showed in Eq. 5, there are two main factors that affect the channel quality: transmission range and the path loss exponent γ. For a predefined network location, the path loss exponent γ is known and can be calculated. Hence, only transmission range can affect the transmission rate. There is a direct relation between communication rate and transmission range following the Eq. 5. The communication rate must be adjusted to the relative distance of a link between two nodes.

In this paper, to support the selection of data rate r_k, MAC layer delivers received data packets to the network layer along with the RSSI for the packet. The RSSI provides the information about receiver power P_r. The received power P_r is used to compare with the referenced sensitivity $P_{S_{r_k}}$ as showed in Table 1. If $P_r \geq P_{S_{r_k}}$, the highest possible rate r_k with the consideration of link reliability is chosen for data transmission. For example, if a node receives a packet with $P_r = -68$ dBm, then it determines $P_{S_{r_k}} = 36$ Mbps because $P_{S_{r_k}}(36 \text{ Mbps}) < -68 \text{ dBm} < P_{S_{r_k}}(48 \text{ Mbps})$. Hence, the highest supported rate in this case is 36 Mbps. The connectivity is broken when the relative distance is greater than $R_{\min(r_k)}$ (the two nodes out of communication range).

3.2 Proposed routing metric

Consider a multi-rate ad hoc network in which any two neighboring nodes (direct communication) use the highest possible rate to communicate corresponding to their relative distance as described in Section 3.1. In the wireless environment, due to the impact of many factors such as interferences and collisions, wireless links are not completely reliable. Hence, a packet may need to be transmitted more than one time in order to be successfully received.

Definition 1 Let d_f and d_r denote the packet delivery ratio in the forward and reverse directions, respectively. Let $\delta_{ab}^{(r_k)}$ be the *link reliability* when node a and node b communicate at rate r_k. Hence, the *link reliability* $\delta_{ab}^{(r_k)}$ is the fraction of packets which are successfully received and can be defined as

$$0 < \delta_{ab}^{(r_k)} = d_f \times d_r \leq 1 \tag{6}$$

Next, consider node i-th belonging to a route from source to destination and define a weight for that position in the route. For a route, the weight associated with the i-th position is the sum of the link weight between the node that occupies the i-th position and the nodes occupying the previous and next positions in the route. Therefore, the weight associated with the i-th position can be defined as:

$$W_i = \frac{\delta_{i-1}^{(r_k)} r_k + \delta_{i+1}^{(r_l)} r_l}{\ln\left(r_{\max} - \min(\delta_{i-1}^{(r_k)} r_k, \delta_{i+1}^{(r_l)} r_l) + e\right)} \tag{7}$$

where r_k and r_l are the maximum possible rates that the node occupying the i-th position and the nodes occupying the previous and next positions can use to communicate, respectively. As mentioned in Section 1, $\delta_{i-1}^{(r_k)}$ and $\delta_{i+1}^{(r_l)}$ denote the link reliability with the previous and next hop of node i-th at rate r_k and r_l, respectively. Nodes with low link rate will have lower weight as defined in the denominator of Eq. 7. And $\delta^{(r_k)} \times r_k$ is the effective link capacity at rate r_k. In this equation, $r_{\max}$ is the maximum data rate that a specific standard supports (in this case $r_{\max} = 54$ Mbps). We use natural logarithm (base $e = 2.718$) for calculation in this paper, although any other logarithm base is acceptable.

If the value of r_k is much different with r_l, for example $r_k \ll r_l$, the route will have link bottleneck between $(i-1)$-th and i-th positions. To avoid choosing that node, we define the cost of node i-th as

$$C_i = \frac{W_i}{\ln\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l + e\right|} \tag{8}$$

The denominator in Eq. 8 is $\ln\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l + e\right|$ to ensure that the value of C_i is finite when $\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l\right| = 0$.

Let N_i be the number of intermediate nodes in the route, and the coefficient α_i of position i-th is defined as

$$\alpha_i = \frac{C_i}{\sum\limits_{i=1}^{N_i} C_i} \tag{9}$$

in which $\sum\limits_{i=1}^{N_i} \alpha_i = 1$. Note that α_i always greater than 0 because for a valid route, there exists at least one link in that route (in case $N_i = 0$ as source and destination are neighbors), and $\alpha_0 = 1$ in this special case.

Finally, the RAI is defined to choose the best route between source/destination pairs

$$\mathrm{RAI} = \begin{cases} \dfrac{1}{N_i}\sum\limits_{i=1}^{N_i} \alpha_i \ln \alpha_i - \ln \dfrac{\min\left(\delta_i^{(r_k)} r_k\right)}{r_{\max}} & \text{if } N_i > 1, \\ -\dfrac{\ln e}{e} - \ln \dfrac{\min\left(\delta_i^{(r_k)} r_k\right)}{r_{\max}} & \text{if } N_i = 0, 1 \end{cases} \tag{10}$$

The metric is obtained as follows: The first part $(\frac{1}{N_i}\sum\limits_{i=1}^{N_i} \alpha_i \ln \alpha_i \leq 0)$ is formulated by applying the entropy properties[1] [7] as an indication of the link's homogeneity and the path length: Route with the minimum difference on capacities among links, and small number of hops, is preferred. The second part $(\ln \frac{\min(\delta_i^{(r_k)} r_k)}{r_{\max}})$ is the reference to choose the high data rate of individual links in a route: Route with high capacity in each node is preferred. Hence, the route with the smaller value of RAI will be proven to be a high-throughput route without bottleneck links and has the small number of hops. The proofs of those properties can be found in the "Appendix."

We will discuss the effect of each parameter in Eq. 10 on finding the best route by the following properties:

Lemma 1 *The value of RAI satisfies the following condition:*

$$-\frac{\ln N_i}{N_i} - \ln \frac{\min\left(\delta_i^{(r_k)} r_k\right)}{r_{\max}} \leq \mathrm{RAI} \leq -\ln \frac{\min\left(\delta_i^{(r_k)} r_k\right)}{r_{\max}}$$

Proof First, in Eq. 9, we have $0 < \alpha_i \leq 1$; hence, $\ln \alpha_i \leq 0$ and $\max \sum\limits_{i=1}^{N_i} \alpha_i \ln \alpha_i = 0$ lead to $\mathrm{RAI} \leq -\ln \frac{\min(\delta_i^{(r_k)} r_k)}{r_{\max}}$.

Next, for a sequence of non-negative number $\{\alpha_i\}$, using *log sum inequality* [7], we have

$$\frac{1}{N_i}\sum_{i=0}^{N_i} \alpha_i \ln \alpha_i \geq \frac{1}{N_i} \ln \sum_{N_i} \sum_{N_i} \alpha_i = -\frac{\ln N_i}{N_i}$$

Hence, $\min \frac{1}{N_i}\sum\limits_{i=1}^{N_i} \alpha_i \ln \alpha_i = -\frac{\ln N_i}{N_i}$.

An alternative way to prove this lemma: Note that $-\sum\limits_{i=1}^{N_i} \alpha_i \ln \alpha_i$ is the entropy function, say $H(X)$ and $H(X) \geq 0$. Thus, the left-hand equality holds. Furthermore, $H(X)$ is maximum when $\alpha_i = 1/N_i$, and in this case, $H(X) = lnN_i$. From this, the right-hand equality holds. The proof details can be found in the "Appendix." Hence, the lemma is proven. □

[1]The entropy's properties for some sets $\{\alpha_i\}$ satisfied $\sum \alpha_i = 1$ can be expressed as follows: (1) *maximality*: Among observing sets which contain the same number of elements in each set, the set with more resemble elements will have the higher entropy outcome. Especially, a set of homogeneous elements will have the maximum entropy outcome. (2) *Uniform distribution*: For sets with the different number of homogeneous elements, the higher number of elements a set has, the lower entropy outcome per element that set gets.

The problem of selecting the best route becomes the problem of choosing a route with the minimum RAI value, as equivalent to maximum cost C_i in each intermediate node. Hence, the following properties of a route are achieved when it has small/minimum RAI value:

Theorem 1 *The route with minimum RAI value, defined by Eq. 10, is the route with the small number of hops between source and destination.*

Proof From Lemma 1, $\mathrm{RAI}_{\min} = -\frac{\ln N_i}{N_i} - \ln \frac{\min(\delta_i^{(r_k)} r_k)}{r_{\max}}$. Therefore, when the number of intermediate nodes increase, we have

$$\lim_{N_i \to \infty} \frac{\ln N_i}{N_i} = 0$$

This means for route 1 with $N_i^{(1)} < N_i^{(2)}$ of route 2, $\mathrm{RAI}^{(1)} < \mathrm{RAI}^{(2)}$ and the value of RAI will be increased as the number of intermediate nodes increases. Moreover, RAI is minimized when α_i has maximum value. From Eq. 9, α_i has higher value when N_i is small and $\max(\alpha_i) = 1$ when $N_i = 0$ (no intermediate node) or $N_i = 1$. Therefore, with N_i small, the value of RAI will be decreased. Hence, the theorem is proven. □

Theorem 2 *The route with minimum RAI value defined by Eq. 10 (consequently has the maximum cost C_i defined by Eq. 8) can avoid link's bottleneck.*

Proof From Lemma 1, the equality holds ($\mathrm{RAI}_{\min} = -\frac{\ln N_i}{N_i} - \ln \frac{\min(\delta_i^{(r_k)} r_k)}{r_{\max}}$). This implies that α_i (or C_i) of any node in the route must have the equal value. Also, maximize the cost C_i in Eq. 8 becomes maximize the weight W_i and minimize $\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l + e\right|$. It means intermediate nodes with high data rate in all two links (minimize $r_{\max} - \min(\delta_{i-1}^{(r_k)} r_k, \delta_{i+1}^{(r_l)} r_l)$) and small difference between two links capacities (minimize $\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l\right|$) are preferred. Hence, the best case is $\delta_{i-1}^{(r_k)} r_k = \delta_{i+1}^{(r_l)} r_l$ (node i-th communicates with the previous and next nodes at the same highest effective link capacity). Contrarily, if there exist some links that have data rate much lower than other links in the route, the RAI value will be increased and that route will not be chosen. □

Theorem 3 *The route with the minimum RAI value, defined by Eq. 10, has the highest throughput among route candidates.*

Proof As showed in Theorem 2, minimize the RAI value is equivalent to maximize C_i. To maximize the value of C_i, the value of W_i in the Eq. 7 needs to be maximized along with minimize the denominator. The value W_i is maximized when $(\delta_{i-1}^{(r_k)} r_k + \delta_{i+1}^{(r_l)} r_l)$ is maximized and $[r_{\max} - \min(\delta_{i-1}^{(r_k)} r_k, \delta_{i+1}^{(r_l)} r_l)]$ is minimized. Hence, an intermediate node will choose the next hop which has the highest effective link capacity with that intermediate node to communicate. Also, the denominator is ensured to be minimized when that node has a small difference between two links capacities as shown in Theorem 2. The process is the same for all links in the route. Also, the length of routes is taken into account as discussed above. Therefore, the selected route will have the highest throughput among route candidates. □

In summary, as the above theorems have showed, the problem of selecting the best route becomes the problem of choosing the route with minimum RAI value (that leads to maximum cost C_i in each intermediate node) among the route candidates. The RAI values of route candidates are directly calculated and compared in the route discovery period. The details will be shown in the next section.

3.3 Route discovery strategy

It is conceivable that any routing metric based on the weight/cost of node/link should only be ensured to find the best path if and only if each node knows the entire network topology. Hence, it is straightforward for those metrics to use in the proactive routing protocol. However, for the reactive routing protocol, since traditional method is to keep track on only one-hop neighbor information, the first received RREQ at the destination apparently cannot guarantee that the found route is either the best or the better route compared with the others. In theory, if the network information is unknown, for example in case of using MTM for reactive routing, every intermediate node in the network needs to forward any copy of the received RREQ to help the destination calculate all possible combinations of metric values. Also, the destination must wait for an unknown time for receiving those copies of a RREQ to find the best route. However, the authors in [3] and [9] did not mention about any specific mechanism for efficient route discovery.

Like MTM or any other weight/cost-based metrics, RAI metric must allow duplicated RREQs retransmitting to find a better route within an acceptable time in route discovery phase. In this section, a simple but effective technique is used to find a high-throughput route. From the motivation that the prior knowledge about neighbors information can support to find the better route (in case of partial knowledge of network topology) or the best one (in case of full network topology information), we propose to use two-hop neighbor information for effective route discovery phase in the Algorithm 1

Algorithm 1 Pseudo-code for 2-hop information maintenance

For a node in the network does
  locally calculates link weights with its neighbors,
  periodically broadcasts the list of its neighbors along with their link weights
  if receives a Hello message from its neighbor,
  then updates and maintains the 2-hop information.
  if receives a RREQ from its neighbor,
  then executes the Alg.2 for the route discovery process.

The advantages of using two-hop information will be shown in the example below. In order to keep track of the topology of the two-hop neighbor, a node should include a list of its one-hop neighbors together with their link weights in the hello messages that it periodically broadcasts to its neighbors. By receiving hello messages from every neighbor, a node is able not only to have a complete view of the one-hop neighbor topology but also to know its two-hop neighbors and their connectivity with the one-hop neighborhood. It means that a node should know the links among the two-hop neighbors to have a complete two-hop topology information.

Figure 1 shows how route discovery works with the assistance of two-hop topology information. To make it simple for understanding, we only consider link rate without link reliability in this example. In the figure, the number (such as 12, 24, 36, 54) in each link indicates the data rate of that link. Suppose the first arrived RREQ to destination will go through route $\{S, I_2, I_6, D\}$. Apparently, AODV based on minimum hop count will select that route for communication. However, that route is not a high-throughput one and contains a link bottleneck (I_2, I_6). Therefore, in case of RAI protocol, when node I_6 receives a RREQ message from I_2, it calculates RAI values of all possible routes from source node up to I_6 itself using the link weights that it knows from two-hop topology. By this way, even though node I_6 still does not receive a copy of that RREQ from other neighbors (except from node I_2), it knows that the best route (within two-hop topology knowledge) from source up to current node I_6 must go through node I_4 as $\{S, I_2, I_4, I_6, \ldots\}$. Hence, node I_6 will continue forward that RREQ to its neighbors with the list of nodes which form the most high-throughput route from destination. In this example, node I_6 will send the list $\{S, I_2, I_4, I_6\}$ instead of list $\{S, I_2, I_6\}$ as traditional protocols will do. Every node that has already been included in the list (i.e., node I_4) will simply discard all latter arriving duplicated RREQ to avoid addition overhead of message flooding. Therefore, route discovery using two-hop information is effectively enough and also can balance the trade-off between the overhead of maintaining network information and time-consuming for finding a good route. We also observed that by using two-hop information, the first arrived RREQ packet in almost cases provides the information of the best route. That is because of latter arrived RREQ packets usually go through routes with the considerable number of addition nodes that make the RAI value increase.

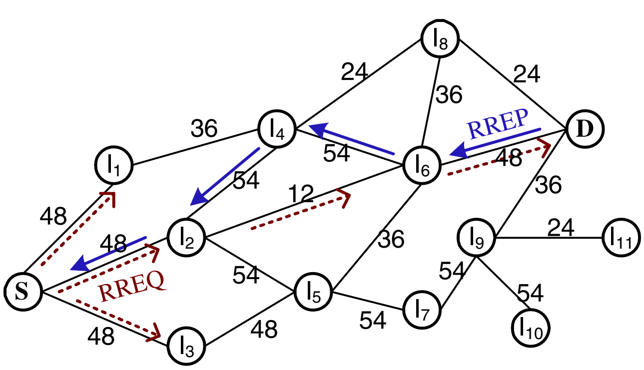


Fig. 1 An example of route discovery

3.4 Protocol operation

Like other on-demand routing protocols, the process of route discovery and maintenance based on the RREQ, RREP, and RERR exchange is used in our protocol. The operation is the same as traditional AODV protocol briefly discussed in Section 2 with some modifications presented below. The source address and sequence number fields in the RREQ jointly identify a unique and instant RREQ in the network. Instead of hop count, we use "Route Assessment Index" as the routing metric. The proposed protocol enables a node to choose and to control the data rate for a packet. The network layer sends a packet to the MAC layer with the desired data rate for a transmission. The MAC cooperates the network layer by delivering the delivery ratio (link reliability) and the received signal strength (or, the RSSI) along with a received packet. From those parameters, the network layer can calculate the link cost and selects the appropriate data rate to adjust with the corresponding distance.

Each node uses a routing table (or cache) for multi-hop communications. The table maintains the route entries in the following format:

{Destination, $\{C_i\}$, N_i};

where $\{C_i\}$ are the list of costs of intermediate nodes from the destination up to current node. For example, in Fig. 1, node I_6 will keep the costs $\{I_2, I_4\}$ because the best route from S up to I_6 based on two-hop information as discussed in Section 3.3 is $\{S, I_2, I_4, I_6\}$. Here, N_i is the cumulative value of the number of intermediate nodes up to that node. A node updates entries in the table if there is any change of those parameters.

Algorithm 2 Pseudo-code for the route discovery algorithm

Initial: each node maintains the 2-hop neighbor information.
For an intermediate node has received $RREQ_i$ does
  if $RREQ_i$ has the same ID with the previous received one,
  then check the list of nodes constructed from the source.
    if the list is similar to the existing list,
    then discards the $RREQ_i$,
    else recalculates RAI value of the route up to that node.
      if the new value is smaller than the current value,
      then updates RAI value, and forwards the $RREQ_i$,
      else discards the $RREQ_i$.
  else calculates RAI values based on its 2-hop information,
      selects the route with the smallest RAI value,
      forwards the $RREQ_i$ with that RAI value and the list of nodes constructed from the source.
repeat until the $RREQ_i$ reaches the destination.

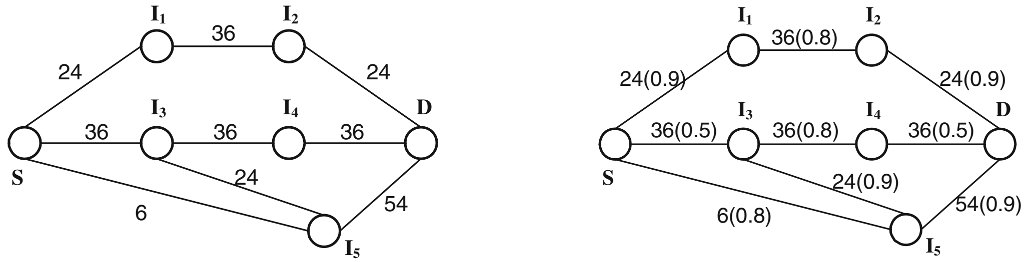
The Algorithm 2 shows the process of route discovery. When a node receives a RREQ, it locally calculates RAI values as described in Section 3.3 and chooses the best one within two-hop information knowledge. Then the RREQ is forwarded with the list of nodes which forms the best route and their costs if the node receives the request for the first time. Unlike AODV which will discard the duplicated RREQ, in our protocol, when a RREQ with the same ID to the previous RREQ arrives a node, that node checks the list of nodes constructed from source node. If the list is different, then it calculates and checks the value of new RAI. If the new value of RAI is lower than the previous value, then it will update that value and forward that RREQ copy. Otherwise, it will discard that RREQ. The process is repeated until a specific RREQ reaches the destination. The destination then calculates the value of α_i and RAI based on information of received RREQ. The first received request at the destination is immediately replied with a unicast RREP packet that contains the RAI value of the route. That route is always the best route within two-hop neighbor information. By this way, the destination does not have to wait for an amount of time to get the better route. If destination receives another RREQ through a better route later (lower RAI value), it overrides the previous route by sending a new RREP. Otherwise, it will discard that duplicated RREQ because the route with higher RAI value will not be chosen. Each time the source node receives an update RREP with lower RAI value, it will use the updated route for delivering data.

Discussion Because the RAI value is calculated locally and only when needed if a node receives a RREQ, the route discovery phase and the computation overhead are negligible. Indeed, for the route discovery phase, the duplicated RREQs are allowed but only when the higher value of RAI needs to be updated. Also, RREQs through lower throughput routes toward a destination are excluded at intermediate nodes using the proposed route discovery strategy. Suppose we have the network with n nodes. By using two-hop information and the above route discovery, the number of broadcast RREQs has the upper bound $O(2n)$. Therefore, the time delay for the route discovery is noticeably reduced. The details will be shown and discussed more in Section 4 below.

For the computation overhead, the calculation of RAI values is distributed among intermediate nodes up to destination to find the local optimal value. Hence, the protocol is distributed orientation. Further more, each node keeps track of the necessary information by passive hearing the hello messages periodically broadcasted from its neighbors. Therefore, the message exchanges of this routing protocol are similar to those of the traditional AODV protocol. In fact, the necessary information only needs to be attached in the periodically broadcasted hello messages, and the RSSI information can be directly measured from those received hello packets to decide the communication rate. The computation complexity is trivial and identical for every intermediate node because each node only needs to calculate RAI values from the source up to that node based on the received information from its neighbors. Hence, the proposed routing protocol is implementable without the concern of computational complexity.

3.5 An example of route selection

We use the network topology with seven nodes as shown in Fig. 2 to demonstrate the proposed protocol operation. Suppose node S needs to find a best route to node D. There are five intermediate nodes $I_1, I_2, \ldots, I_5$ with communication links and data rates corresponding to their relative distances as shown in the figure. There are five possible routes from S to D, which are (1) $\{S, I_1, I_2, D\}$, (2) $\{S, I_3, I_4, D\}$, (3)



(a) Case I: Without the impact of link reliability ($\delta_{ab}^{(r_k)} = 1$), the best route is $\{S, I_3, I_4, D\}$

(b) Case II: With the impact of link reliability ($0 \leq \delta_{ab}^{(r_k)} < 1$), the best route is $\{S, I_1, I_2, D\}$

Fig. 2 An example of route selection using RAI metric. **a** Case I: without the impact of link reliability ($\delta_{ab}^{(r_k)} = 1$), the best route is $\{S, I_3, I_4, D\}$. **b** Case II: with the impact of link reliability ($0 \leq \delta_{ab}^{(r_k)} < 1$), the best route is $\{S, I_1, I_2, D\}$

$\{S, I_3, I_5, D\}$, (4) $\{S, I_5, D\}$, and (5) $\{S, I_5, I_3, I_4, D\}$, respectively. To demonstrate how does RAI metric work and what happen if the link reliability is taken into account, we consider two cases: link without packet loss (link reliability $\delta_{ab}^{(r_k)} = 1$) and link with packet loss ($0 \leq \delta_{ab}^{(r_k)} < 1$).

3.5.1 Case I: without impact of link reliability

After selecting the communication rate based on the received RSSI signal, as shown in Fig. 2a, the RAI metric is used to calculate RAI value of each route as mentioned in Section 3.4. In the figure, the number (such as 6, 24, 36, 54) in each link indicates the data rate of that link.

- *For route* 1: From Eq. 8 and Eq. 9, we have $C_1^{(1)} = C_2^{(1)} = 6.39$ and $\alpha_1^{(1)} = \alpha_2^{(1)} = 0.5$, respectively. Hence, from Eq. 10, we have $\text{RAI}^{(1)} = 0.46$.

Similarly, we have

- *For route* 2: $C_1^{(2)} = C_2^{(2)} = 23.7$, $\alpha_1^{(2)} = \alpha_2^{(2)} = 0.5$, and $\text{RAI}^{(2)} = 0.06$.
- *For route* 3: $C_1^{(3)} = 6.39$, $C_2^{(3)} = 6.41$, $\alpha_1^{(3)} = 0.499$, $\alpha_2^{(3)} = 0.501$, and $\text{RAI}^{(3)} = 0.464$.
- *For route* 4: $C_1^{(4)} = 3.89$, $\alpha_1^{(4)} = 1$, and $\text{RAI}^{(4)} = 1.829$.
- *For route* 5: As discussed in Section 3.3, if RREQ arrives to either node I_3 or I_5 first, those nodes know the link cost of each other. Therefore, node I_3 will know the best route from S to it is $\{S, I_3\}$ and node I_5 will know the best route from S to it is $\{S, I_3, I_5\}$. For that reason, the route $\{S, I_5, I_3\}$ from S to I_3 will not be considered, and the destination does not have to calculate the value of $\text{RAI}^{(5)}$. Thanks to two-hop neighbor information, the lower throughput routes will be excluded at intermediate nodes and the arrived RREQs always carry the information of the better routes. If we calculate the RAI value of route 5 to verify the metric's correctness, we have $C_1^{(5)} = 2.52$, $C_2^{(5)} = 6.39$, $C_3^{(5)} = 23.7$, $\alpha_1^{(5)} = 0.077$, $\alpha_2^{(5)} = 0.196$, $\alpha_3^{(5)} = 0.727$, and $\text{RAI}^{(5)} = 1.948$.

Therefore, $\text{RAI}^{(2)} < \text{RAI}^{(1)} < \text{RAI}^{(3)} < \text{RAI}^{(4)} < \text{RAI}^{(5)}$; hence, the best route is route 2 with $\text{RAI}^{(2)} = 0.06$.

3.5.2 Case II: with impact of link reliability

In this case, each link has a link reliability denoted along with link rate as shown in Fig. 2b. Similarly, we can calculate the RAI value of each route as follows:

- *For route* 1: $C_1^{(1)} = C_2^{(1)} = 6.137$, $\alpha_1^{(1)} = \alpha_2^{(1)} = 0.5$, and $\text{RAI}^{(1)} = 0.57$.
- *For route* 2: $C_1^{(2)} = C_2^{(2)} = 4.92$, $\alpha_1^{(2)} = \alpha_2^{(2)} = 0.5$, and $\text{RAI}^{(2)} = 0.75$.
- *For route* 3: $C_1^{(3)} = 6.173$, $C_2^{(3)} = 5.816$, $\alpha_1^{(3)} = 0.515$, $\alpha_2^{(3)} = 0.485$, and $\text{RAI}^{(3)} = 0.752$.
- *For route* 4: $C_1^{(4)} = 3.521$, $\alpha_1^{(4)} = 1$, and $\text{RAI}^{(4)} = 2.052$.
- *For route* 5: Similar to case I, route 5 is actually not considered because the best route to I_3 from S is still a direct link (S, I_3).

Therefore, $\text{RAI}^{(1)} < \text{RAI}^{(2)} < \text{RAI}^{(3)} < \text{RAI}^{(4)}$; hence, the best route is route 1 with $\text{RAI}^{(1)} = 0.57$.

Observations In the case I, when link reliability is not considered, the routing protocol will choose the route with high rate in each link and avoid link bottleneck. However, when the link reliability is taken into account in the case II, the bottleneck free route with actually high-throughput links (the highest effective link capacity) is chosen. Also, in this example, we can see that even route 4 has the smallest intermediate node (node I_5); in both case I and case II, that route had not been chosen because it contains link at lowest rate (6 Mbps) that makes the RAI value increase. Therefore, the end-to-end throughput of route 4 is significantly downgraded because of that bottleneck link. The proposed protocol with link reliability reflects the real situation in its routing metric, hence can accurately select the best route.

4 Performance analysis

We evaluate the performance of proposed multi-rate routing metric named RAI using *NS-2* [1] to compare with the traditional *hop-count* metric of AODV [16], WCETT metric [9], and MTM [3]. The original WCETT was implemented with both single and multiple radios. However, in this paper, we consider only single wireless interface for all routing metrics. Therefore, the channel diversity is not considered in WCETT. The network with the number of nodes varying from 50 to 250 is randomly distributed over a 500 × 500-m area. Each node can send/receive data packets at any of the IEEE 802.11g-supported data rates (i.e., 6, 9, 12, 18, 24, 36, 48, or 54 Mbps) and uses IEEE 802.11 DCF for channel access. According to the relative distance with its neighbors, a node will use the highest possible rate to communicate following the Eq. 5. Only for the study case of the impact of traffic load, we manually inject the traffic flow with different data rates and different number of simultaneous flows. The setup details are given in each simulation scenario presented below. We pick up some source–destination pairs randomly. UDP flows with the packet size set to 1,024 bytes are applied in the source nodes. The log-distance path loss radio propagation model discussed in Section 3.1 is used with the path loss exponent $\gamma = 3$. Each simulation run has been executed 20 times, and the average results are plotted in the graphs.

We observe average route discovery time of the mentioned metrics above. Figure 3 shows that in all cases, the discovery time increases sharply when the number of hops (path length) between source and destination increases. This is because the more intermediate nodes in the path, the more medium access contentions will occur that cause more time consumption. AODV allows only one RREQ per node to find the minimum hop route. Each node is expected to forward the RREQ only once (totally $O(n)$ broadcasts with n number of nodes); hence, apparently its discovery time is the shortest. In cases of WECTT and MTM, theoretically, the metrics are efficient in selecting the optimum route. However, only forwarding the first RREQ, which is used in the original AODV, does not guarantee that the RREQ for the optimum path will be forwarded.

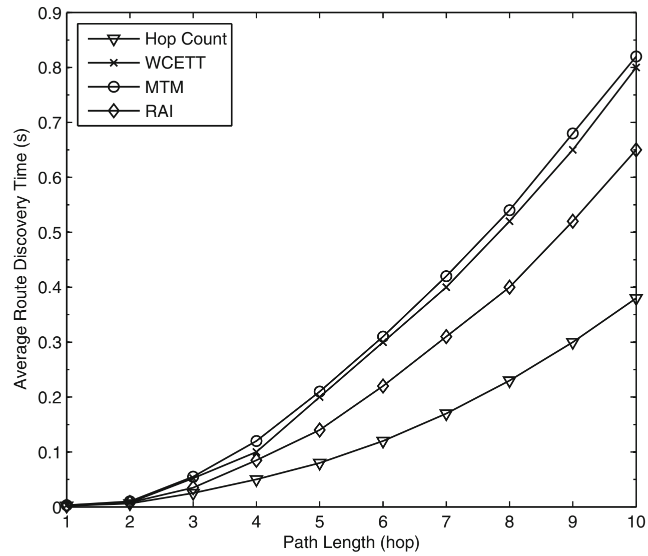


Fig. 3 Average route discovery time

The destination can select the optimal route only when it receives all possible combinations. Hence, all intermediate hops in the network need to forward every copy of the received RREQ (requires $O(n^2)$ RREQ broadcasts). Because of that, WECTT and MTM discovery times are almost same. For RAI, duplicated RREQs are allowed but only when the higher value of RAI needs to be updated. Also, RREQs through lower throughput routes toward destination are excluded at intermediate nodes using the above discussed route discovery strategy. Therefore, the maximum time delay in the worst case is $O(2n)$ because a node needs maximum two times to forward a RREQ for finding the best value within the two-hop information. Hence, the delay time is much smaller than both WCETT and MTM which need to get all possible combinations of their values before selecting a route.

Second, we evaluate the average end-to-end throughput for different path lengths. We randomly generate ten simultaneous flows with various distances from 20 to 500 m. We can observe some interesting results here. When the distance is far, more hops need to be used to reach the destination. We can observe that the end-to-end throughput of a short route (need a few hops to reach destination) is degraded sharply for each addition hop. For example, the throughput of 50 m length route is almost less than a half of the throughput of 20 m length route because it may need one more hop to reach the destination. However, the throughput deduction is less severely for longer route, and the throughput seems to plateau. For example, as shown in Fig. 4, for RAI metric, the end-to-end throughput of 400 m length route is around

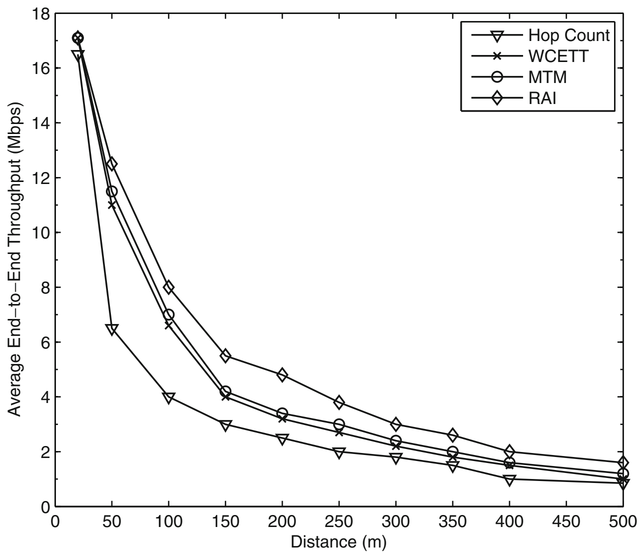


Fig. 4 Average end-to-end throughput with varying distances

2 Mbps while that of 500 m length route is still around 1.6 Mbps. Those phenomena can be explained as the following reasons: When packets are sent along a route in a multi-hop network, the adjacent transmissions are competing for access to medium. Also, to avoid collision, all nodes within the interference range of the transmitting node must stay idle. When the route is short, for example routes within five-hop length, all other nodes in that route must defer from sending when the node at the third position is transmitting. As the route becomes longer, multiple transmissions can take place simultaneously along the path due to spatial reuse. This allows the throughput to reach a steady state, where an additional distance does not cause any significant decrease in throughput.

Back to the results in Fig. 4, at any distance, RAI performs better than the others with the improvement of about 5% to 88% depending on specific distance. When the path length becomes longer, even though the throughput is downgraded rapidly, the deduction of RAI throughput is less severe than the remaining observing metrics. For MTM, it uses path with the shortest deliver time; hence, it is better than *hop count*. However, without an efficient route discovery strategy, the selected route for data delivery in case of WCETT and MTM metrics in almost case is not the best one. Also, for longer distance with more hops between end nodes, MTM does not consider the route with small value of $\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l\right|$ at all intermediate nodes like RAI does. Hence, the drop rate will be high at bottleneck links, and it will downgrade the throughput. In case of WCETT, it did not consider the back off delay in each transmission. Also, we do not consider the channel diversity for WCETT in this paper; hence, it performs a little bit worse than MTM.

Third, we study the impact of traffic load on the network throughput, as the illustration of the traffic demand variation during the network operation in reality. To do that, we have two alternative ways to observe the impact of traffic load on the routing metrics:

- Gradually increase the packet injection rate of a fixed number of flows.
- Fix the packet injection rate and increase the number of simultaneous flows.

For the first scenario, we setup the network with 100 nodes and adjust the packet injection rate of ten distributed and simultaneous flows, with the packet size 1,024 bytes, from 10 to 50 pkts/s. Figure 5 shows that when the traffic load increases, the achieved network throughput of all routing metrics is also increased. However, it does not increase at the expected rate with the increasing loads. Indeed, network congestions and collisions occur more frequently when the network has simultaneous flows with high traffic loads. The achieved throughput of RAI is higher than that of other observing metrics from 7% to 40% depending on different traffic loads applied to different metrics.

For the second scenario, we setup the network with 100 nodes, fix the packet injection rate to 20 pkts/s, with the packet size 1,024 bytes, and adjust the number of simultaneous flows from five to 25 flows. The results are shown in Fig. 6. From the results, we can observe that with a fixed number of nodes in the network, the net-

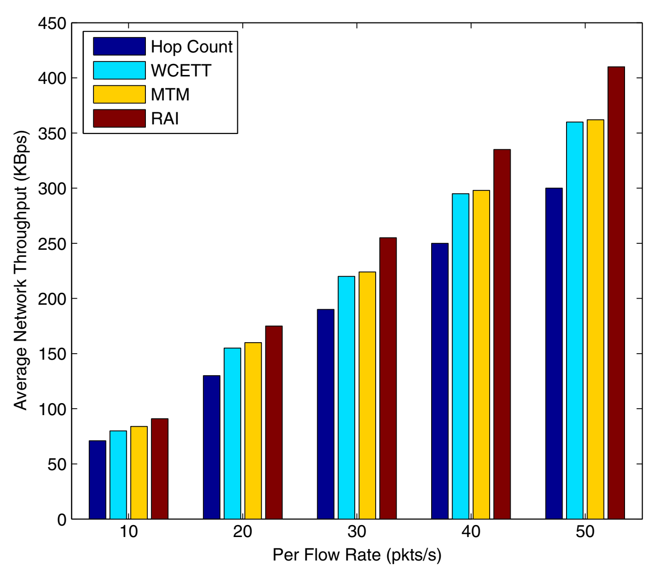


Fig. 5 Impact of traffic load with varying per flow rates

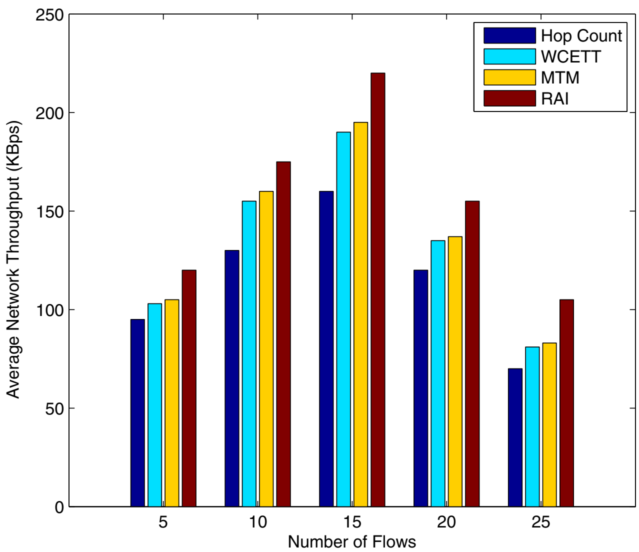


Fig. 6 Impact of traffic load with varying simultaneous flows

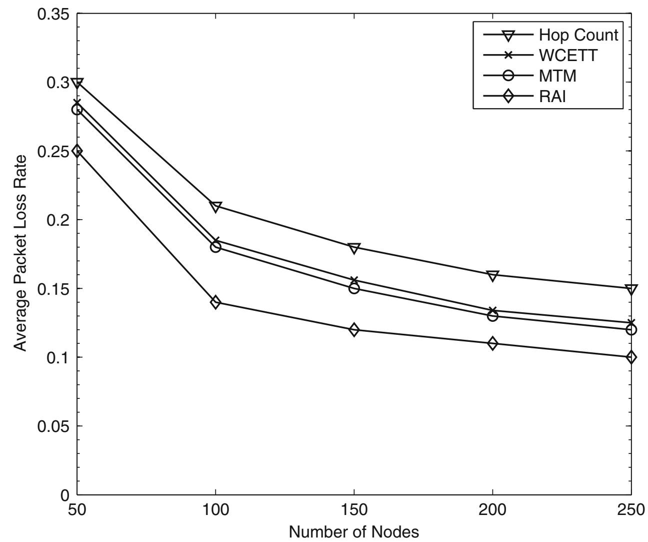


Fig. 7 Average packet loss rate

work throughput increases when the traffic load (traffic per flow, or traffic from simultaneous flows) increases up to an upper bound (which varies following network density and other network conditions). Then, when the number of simultaneous flows continue to increase, the network throughput is decreased because of congested traffic demands. The throughput is even downgraded more seriously when the number of source/destination nodes reaches half of the total nodes in the network. Because there are a lot of contentions, the back off time and the defer time at each node are seriously increased. Only RAI metric designed to avoid link bottleneck can alleviate the throughput deduction. Therefore, it has the best performance compared to the others.

We also consider the following performance metrics under the variation of the number of nodes:

1. *Packet loss rate*: the ratio of the packets that are lost in the route to the number of packets generated by the sources
2. *End-to-end delay*: the average delay experienced by all successfully delivered packets
3. *Network throughput*: the sum of the size of the total data packets received by the destinations per unit time

The results show that RAI outperforms *hop count*, WCETT, and MTM for all performance metrics. For the packet loss rate, it is reduced when the number of nodes increases because the network connectivity is high. As mentioned above, RAI selects route with small value of $\left|\delta_{i-1}^{(r_k)} r_k - \delta_{i+1}^{(r_l)} r_l\right|$; hence, it will limit the link's bottleneck and packet loss rate due to buffer overflows. As shown in Fig. 7, RAI limits the loss rate better than the others.

For the end-to-end delay, in Fig. 8, RAI can reduce from 35% to 55% delay time compared to the others. The reason is that RAI chooses the route with high throughput and less number of intermediate nodes. Also, the link reliability is included; hence, the actual capacity of a link is considered. Hence, the end-to-end delay is the smallest among those observing metrics.

For the average network throughput, it is also increased when the network density is high because when the number of nodes increases, there are more chances

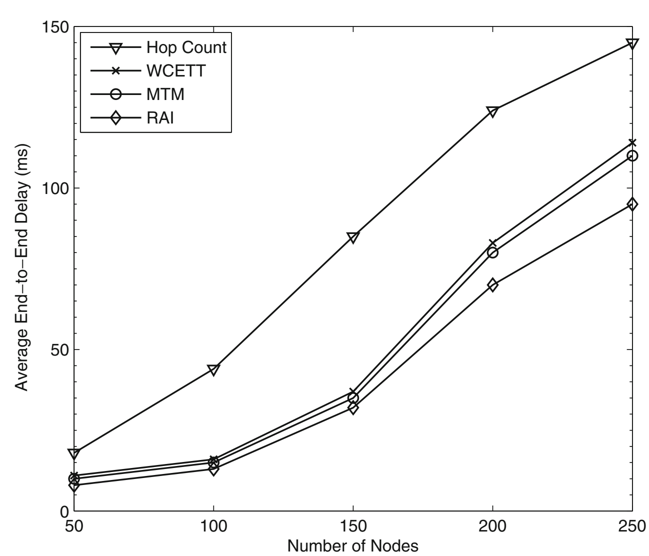


Fig. 8 Average end-to-end delay

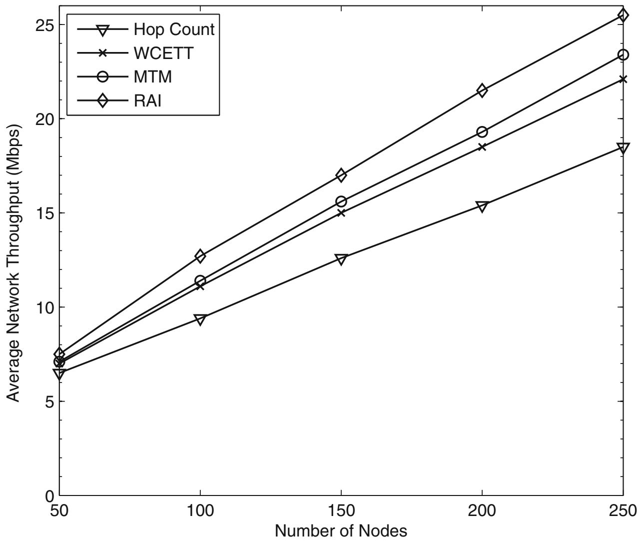


Fig. 9 Average network throughput

for routing protocol to find a better path. In Fig. 9, RAI performs better than *hop count*, WCETT, and MTM with the improvement of about 18% to 40% depending on the network density. RAI, by using appropriate data rate under the effects of network conditions, can improve the network throughput. Indeed, the received RSSI reflects network conditions, such as the path loss exponent γ, inter-flow interference (the interference suffered among concurrent flows), and intra-flow interference (occurs when nodes in a single path attempt to transmit packets of the same flow and interfere with each other), through its value. Also, RAI metric itself considers the link reliability when calculating its value. Hence, the RAI metric effectively chooses the best route under the real network conditions to get better performance of both end-to-end throughput and network throughput.

5 Concluding remarks

As showed in many existing literatures, routing protocols with the supported information from lower layer can perform better because they take into account the actual conditions of the networks. In this paper, we proposed a new routing protocol based on the RSSI and link reliability, which reflect actual network conditions and decide the communication rate of a links, for each node. The proposed routing metric supports reliable and high-throughput route selection for multi-rate ad hoc networks. The route with link bottleneck free and small relay hops is also preferred by minimizing the value of *Route Assessment Index*. Along with the proposed metric, a route discovery strategy is introduced to find an optimal route within an inconsiderable delay. The corresponding proofs and simulation results have showed that the proposed metric with effective route discovery performs significantly better than the existing routing metrics. The proposed routing protocol is implementable and can be applied for different multi-rate ad hoc network standards, which offer reliable and high-throughput applications.

Acknowledgements This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) "(NIPA-2011-C1090-1131-0005). Dr. CS Hong is the corresponding author."

Appendix: Properties of entropy function applied for the route assessment

Let X be a discrete random variable with alphabet χ and probability mass function $p(x) = \Pr\{X = x\}, x \in \chi$. The discrete random variables have a finite number of possible values $(x_1, x_2, \ldots, x_n)$ with probabilities $(p_1, p_2, \ldots, p_n)$, respectively, such that $p_i \geq 0, i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} p_i = 1$. The entropy of that set is defined as

$$H(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} = -\sum_{i=1}^{n} p_i \log p_i \tag{11}$$

The function $H(p_1, p_2, \ldots, p_n)$ (or $H(P)$) is a non-negative, continuous, and symmetric function [7]. In the paper, the observing sets are the set of the coefficient $\{\alpha_i\}$ for route candidates instead of $\{p_i\}$.

Property 1 (Maximality)
The entropy $H(p_1, p_2, \ldots, p_n)$ is maximum when all the probabilities are equal.

$$H(P) = H(p_1, p_2, \ldots, p_n) \leq H\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right) = \log n, \tag{12}$$

with equality if and only if $p_i = 1/n, \forall i = 1, 2, \ldots, n$.

Proof The proof is straightforward following the information inequality theorem in [7] as follows: The relative entropy of two distributed function satisfies

$$D(p \parallel q) = \sum_{x \in \chi} p(x) \log \frac{p(x)}{q(x)} \geq 0 \tag{13}$$

with equality if and only if $p(x) = q(x)$ for all x. Also, from Eq. 11, we have

$$H\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right) = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = \log n$$

Applying the Eq. 13 for two distributed function $H(P)$ and $H(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$, we have

$$D\left(H(P) \parallel H\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)\right) = \log n - H(P) \geq 0$$

Therefore, $H(P) \leq \log n$, the maximality properties is proven. □

It also can be shown that in the general case, for two sets $(p_1, p_2, \ldots, p_n)$ and $(q_1, q_2, \ldots, q_n)$ that satisfy $q_1 \geq q_2 \geq \ldots \geq q_n$, also $p_1 = q_1 + \Delta_u$, $p_2 = q_2 - \Delta_u$, and $p_i = q_i, \forall i = 3, 4 \ldots, n$, we have:

$$H(p) \leq H(q) \tag{14}$$

To prove this property, we have a lemma as follows:

Lemma 2 (Recursive property) *For $n \geq 3$, we have*

$$H(P) = H(p_1 + p_2, p_3, \ldots, p_n) + (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) \tag{15}$$

Proof The proof is straightforward and had been shown in [7]. □

Using Lemma 2, we have

$$\begin{aligned} H(P) &= H(p_1 + p_2, p_3, \ldots, p_n) \\ &\quad + (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) \\ &= H(q_1 + q_2, q_3, \ldots, q_n) \\ &\quad + (q_1 + q_2) H\left(\frac{q_1 + \Delta_u}{q_1 + q_2}, \frac{q_2 - \Delta_u}{q_1 + q_2}\right), \end{aligned} \tag{16}$$

and

$$H(Q) = H(q_1 + q_2, q_3, \ldots, q_n) + (q_1 + q_2) H\left(\frac{q_1}{q_1 + q_2}, \frac{q_2}{q_1 + q_2}\right) \tag{17}$$

Therefore, from Eqs. 16 and 17, we need to prove that $H\left(\frac{q_1+\Delta_u}{q_1+q_2}, \frac{q_2-\Delta_u}{q_1+q_2}\right) \leq H\left(\frac{q_1}{q_1+q_2}, \frac{q_2}{q_1+q_2}\right)$.

Denote $x \triangleq \frac{q_1}{q_1+q_2}$, then $x \geq 0.5$ since $q_1 \geq q_2$. We have

$$\begin{aligned} & H\left(\frac{q_1}{q_1 + q_2}, \frac{q_2}{q_1 + q_2}\right) \\ &= H(x, 1 - x) \\ &= -x \log(x) - (1 - x) \log(1 - x) \\ &\triangleq f(x) \end{aligned} \tag{18}$$

$f(x)$ is a decreasing function for $0.5 \leq x \leq 1$, and then $f(x) \geq f\left(x + \frac{\Delta_u}{q_1+q_2}\right) = H\left(\frac{q_1+\Delta_u}{q_1+q_2}, \frac{q_2-\Delta_u}{q_1+q_2}\right)$. Hence, the general case is proven. It means that for two sets with the same number of elements in each set, the set with more resemble elements will have the higher entropy outcome.

Property 2 (Uniform Distribution)
Suppose we have $\phi(n) = H(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}), n \geq 2, n \in \mathbf{N}$, then $\frac{\phi(n)}{n} \geq \frac{\phi(n+1)}{n+1}$.

Proof The proof is straightforward using the result of Property 1. We have $\phi(n) = \log n$ and $\phi(n + 1) = \log(n + 1)$. Also, $\lim_{n\to\infty} \frac{\log n}{n} = 0$. Therefore, the entropy function will reduce when the number of elements in the set increases (n large) and the theorem is proven. It means that for sets with the different number of homogeneous elements, the higher number of elements a set has, the lower entropy outcome per element that set gets. □

References

1. The Network Simulator—Ns-2. http://www.isi.edu/nsnam/ns/index.html
2. Awerbuch B, Holmer D, Rubens H (2004) High throughput route selection in multi-rate ad hoc wireless networks. In: WONS, pp 253–270
3. Awerbuch B, Holmer D, Rubens H (2006) The medium time metric: high throughput route selection in multi-rate ad hoc wireless networks. MONET 11(2):253–266
4. Cheng MX, Gong X (2010) Interference-aware multipath routing and link rate control in multihop wireless networks. In: GLOBECOM, pp 1–6
5. Clausen T, Jacquet P (2003) Optimized link state routing protocol (OLSR). Network Working Group, Project Hipercom, INRIA
6. Couto DSJD, Aguayo D, Bicket JC, Morris R (2003) A high-throughput path metric for multi-hop wireless routing. In: MOBICOM, pp 134–146
7. Cover TM, Thomas JA (2006) Elements of information theory, 2nd edn. Wiley-Interscience, Hoboken
8. Draves R, Padhye J, Zill B (2004) Comparison of routing metrics for static multi-hop wireless networks. In: SIGCOMM, pp 133–144

9. Draves R, Padhye J, Zill B (2004) Routing in multi-radio, multi-hop wireless mesh networks. In: MOBICOM, pp 114–128
10. Fan Z (2004) High throughput reactive routing in multi-rate ad hoc networks. Electron Lett 40(25):1591–1592. doi:10.1049/el:20046622
11. Holland G, Vaidya NH, Bahl P (2001) A rate-adaptive MAC protocol for multi-hop wireless networks. In: MOBICOM, pp 236–251
12. Hong K, Sengupta S, Chandramouli R (2010) Cross-layer MAC enabling virtual link for multi-hop routing in wireless ad hoc networks. In: IEEE international conference on Communications (ICC), 2010, pp 1–5. doi:10.1109/ICC.2010.5502371
13. IEEE (2007) IEEE: standard for information technology—telecommunications and information exchange between systems—local and metropolitan area networks—part 11: wireless LAN MAC and physical layer (PHY) specifications. IEEE Std 802.11-2007
14. Johnson DB, Maltz DA, Broch J (2001) DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks. In: Perkins CE (ed) In ad hoc networking, chapter 5. Addison-Wesley, Boston, pp 139–172
15. Kamerman A, Monteban L (1997) Wavelan(C)-II: a high-performance wireless LAN for the unlicensed band. Bell Labs Tech J 2(3):118–133
16. Perkins CE, Belding-Royer EM (1999) Ad-hoc on-demand distance vector routing. In: WMCSA, pp 90–100
17. Perkins CE, Bhagwat P (1994) Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In: SIGCOMM, pp 234–244
18. Rappaport TS (1995) Wireless communications: principles and practice. Prentice Hall PTR, Upper Saddle River
19. Sadeghi B, Kanodia V, Sabharwal A, Knightly EW (2002) Opportunistic media access for multirate ad hoc networks. In: MOBICOM, pp 24–35
20. Shah GA, Liang W, Shen X (2010) Cross-layer design for QoS support in wireless multimedia sensor networks. In: GLOBECOM, pp 1–5
21. Trong H, Lee S, Hong C (2010) End-to-end throughput improvement for single radio multi-channel multi-path wireless mesh networks: a cross layer design. Ann Telecommun 65:635–646. doi:10.1007/s12243-010-0178-y
22. Uddin M, Alazemi H, Assi C (2010) Joint routing, scheduling and variable-width channel allocation for multi-hop WMNS. In: IEEE international conference on Communications (ICC), 2010, pp 1–6. doi:10.1109/ICC.2010.5502362