Wireless Netw (2006) 12:797–812
DOI 10.1007/s11276-006-6056-x

Location-aware resource management in mobile ad hoc networks

Xiang Chen · Wei Liu · Hongqiang Zhai · Yuguang Fang

Published online: 27 April 2006

Abstract We propose an innovative resource management scheme for TDMA based mobile ad hoc networks. Since communications between some important nodes in the network are more critical, they should be accepted by the network with high priority in terms of network resource usage and quality of service (QoS) support. In this scheme, we design a location-aware bandwidth pre-reservation mechanism, which takes advantage of each mobile node's geographic location information to pre-reserve bandwidth for such high priority connections and thus greatly reduces potential scheduling conflicts for transmissions. In addition, an end-to-end bandwidth calculation and reservation algorithm is proposed to make use of the pre-reserved bandwidth. In this way, time slot collisions among different connections and in adjacent wireless links along a connection can be reduced so that more high priority connections can be accepted into the network without seriously hurting admissions of other connections. The salient feature of our scheme is the collaboration between the routing and MAC layer that results in the more efficient spatial reuse of limited resources, which demonstrates how cross-layer design leads to better performance in QoS support. Extensive simulations show that our scheme can successfully provide better communication quality to important nodes at a relatively low price. Finally, several design issues and future work are discussed.

Keywords Mobile ad hoc networks · Bandwidth pre-reservation · Location

X. Chen
Motorola Labs, Schaumburg, IL 60196
e-mail: steven.x.chen@motorola.com

W. Liu
Scalable Network Technologies, Los Angeles, CA 90045
e-mail: liuw@ufl.edu

H. Zhai · Y. Fang (✉)
Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611
e-mail: {zhai; fang}@ece.ufl.edu

1. Introduction

As wireless mobile ad hoc networks (MANETs) are finding more applications in many fields, such as battlefield communications, disaster rescue, and inimical environment monitoring, there is a growing need to provide quality of service (QoS). Due to the lack of fixed wired infrastructure in ad hoc networks, all the communications between wireless mobile nodes are conducted over bandwidth limited and error-prone wireless links. The information exchanges among these nodes may involve multiple hops, with intermediate nodes acting as routers in between sources and destinations. The salient features of MANETs and the application-specific requirements pose great challenges on both the routing layer and the media access control (MAC) layer. Routing protocols must be able to handle route breakage and re-routing, as any failure of the intermediate nodes on the route due to power outage or mobility will disable the entire route. Meanwhile, MAC protocols should effectively reduce access collisions and achieve high channel utilization.

Moreover, a wireless mobile ad hoc network should support service prioritization in terms of network access, given its limited resources.[1] In other words, some important traffic needs to be accepted by the network with higher priority than others. For instance, in a battlefield, the communication requests among some army commanders should be always accepted with higher priority compared with the communications between soldiers since the commanders may need to discuss and coordinate the strategic plans. Another typical scenario is that in cluster based hierarchical ad hoc networks, connections among cluster heads should be specially treated. In such cases, the communications between such important nodes (like commanders, cluster heads, or nodes that take over control tasks in the network, henceforth called *i-nodes*) should be granted higher priority in terms of network access. Thus, the task to support these important communications between these important nodes with QoS guarantee is imperative.

However, in current literature, the research on how to support such important communications with QoS and high acceptance ratio is still thin on the ground at either routing layer or MAC layer. At the routing layer, a variety of QoS routing schemes have been proposed, with the primary goal to find a path from the source to the destination that satisfies the desired QoS requirements [7–11, 19, 21]. In these schemes, whenever there is a connection request, the routing protocol will attempt to find a path with sufficient resources (i.e., bandwidth) to support QoS requirements. If there is no such a path, the request will be rejected. This is undesirable for i-nodes, since in most cases, the connections between them carry important and time-critical information. Moreover, these schemes treat all connections indiscriminately and thus do not appear satisfactory to accommodate the aforementioned important communications between those high-profiled nodes. Obviously, without an effective and efficient resource management scheme,it is very hard for QoS routing schemes to overcome this deficiency.

Meanwhile, many contention-based MAC protocols were proposed to provide differentiated service [1–3, 22]. However, it is known that reserving resources in contention-based MAC is difficult and hence may fail to provide high degree of service guarantee. As a consequence, some slot-based MAC schemes were designed [4–6]. Though they find their niche in wireless local area networks, unfortunately, for multi-hop connections, they alone still cannot support those important communications with QoS.

In fact, the routing layer and the MAC layer should collaborate. Instead of randomly allocating the available resource to the newly accepted connections, the routing layer should intelligently allocate the resource for a connection along its entire path such that the residual resource can accommodate more connections. This means that the routing layer needs to know the allocation restrictions due to MAC collisions of time slots. In addition, a desirable routing protocol also helps the MAC layer maintain high resource utilization. Further, it is observed that a node's location or movement is valuable information that could be utilized to assist both layers in fulfilling the task. All these observations thus motivate us to resort to approaches that are based on the cooperation of these two layers and the use of nodes' location information.

On the other hand, to support the communications between i-nodes, the above mentioned collaborative approach is not adequate. Inspired by the resource reservation concept in cellular networks, we believe it is advantageous to pre-reserve some resources so that *i-connections*, which are defined as connections between any two i-nodes, can have higher acceptance probability. Although many reservation schemes (for example, the *Guard Channel* scheme was proposed in cellular networks to give high priority to handoff calls over newly incoming calls [18]) have been proposed for cellular networks, due to the tremendous difference between MANETs and cellular networks, they cannot be directly applied to our case. As can be imagined, resource reservation is much more complicated in MANETs. Specifically, in cellular networks, bandwidth is reserved at base station, which is static and has good knowledge about the mobility of all the mobile users in the cell it serves. Thus, bandwidth can be reserved easily and effectively; and once reserved, it is always available to mobile users. By contrast, in ad hoc networks, bandwidth is reserved at each mobile node that may move, and most likely is not aware of other nodes' mobility, which greatly increases the difficulty of making efficient resource reservation. Therefore, designing effective and efficient bandwidth pre-reservation in ad hoc networks is not trivial.

In this paper, by introducing the resource pre-reservation[2] into MANETs and taking advantage of nodes' location information, we propose a resource management scheme that is designed on the basis of cooperation between the routing layer and the MAC layer in a TDMA based ad hoc networks. In this scheme, we grant higher priority to connections between any two i-nodes through location-aware bandwidth pre-reservation. By utilizing each node's geographic location information and forming a quadrangle-shaped area between any two i-nodes, the bandwidth pre-reservation mechanism pre-reserves bandwidth for i-nodes along the possible paths connecting any two i-nodes in such a way that potential collisions of time slots are significantly reduced. Subsequently, we use location-aware forward algorithm (LAFA), a bandwidth calculation scheme, which also exploits the geographic information of each node on the path, to find and reserve enough bandwidth to support QoS while keeping the residual resource more efficient for

[1] Unless otherwise specified, resource and bandwidth are interchangeable in this paper.

[2] To distinguish from bandwidth reservation in QoS routing, we call it bandwidth pre-reservation.

future communications. As a result, the network bandwidth is used more efficiently due to the reduction in MAC slot collisions. Therefore, the connection blocking probability for i-connections is greatly decreased without seriously hurting admissions of other connections.

The remainder of this paper is organized as follows. In Section 2, we describe the related work. The system model is introduced in Section 3. Our location-aware bandwidth pre-reservation scheme and the LAFA algorithm are presented in Section 4 and Section 5, respectively. In Section 6, performance evaluation is given with some insightful discussions. We discuss some important issues and future work in Section 7. Finally, concluding remarks are given in the Section 8.

2. Related work

Various schemes have been proposed to address QoS at both routing layers and MAC layer. On MAC protocol design, much work has been done to support QoS. In [1–3, 22], improvements based on random access MAC protocol such as the IEEE 802.11 were proposed. In [1], a distributed approach supporting service differentiation was proposed for wireless packet networks. In [22], Ada and Castelluccia proposed to scale the contention window, use different interframe space or maximum frame length in order to support differentiated services in wireless LANs. Based on the work in [1], Ahn et al. [2] proposed a stateless network model with distributed control algorithm to deliver differentiated service in mobile ad hoc networks, where services are regulated to deal with mobility or traffic overloading. Kanodia et al. [3] combined two mechanisms, i.e., distributed priority scheduling and multi-hop coordination to provide QoS in random access multi-hop wireless networks. Since most of current contention-based MAC schemes cannot provide users with guaranteed QoS, many researchers have looked into the Time Division Multiple Access (TDMA) MAC and proposed various techniques to address issues such as frame length and slot allocation strategies in ad hoc networks [4–6]. Chlamtac et al. [4] proposed a technique called protocol threading to improve the performance of time-spread multiple-access (TSMA) protocols. In [5] and [6], optimization of Time Division Multiple Access (TDMA) frame length and slot allocation strategies were studied in ad hoc networks. Nesargi and Prakash focused on dynamically allocating channels to different cells in cellular mobile networks such that there is no transmission conflict [20]. However, this is different from the ad hoc networks discussed in our paper, where there is no base station.

Considerable effort has also been spent on QoS routing with the primary goal to find a path from the source to the destination that satisfies the desired QoS requirement. In [7] and [8], a scheme to calculate the end-to-end bandwidth of a path under a CDMA-over-TDMA mechanism is proposed. By relaxing the CDMA-over-TDMA MAC scheme, Liao et al. proposed a TDMA-based bandwidth reservation scheme for QoS routing in ad hoc networks [21]. Zhu and Corson [9] proposed an efficient algorithm called forward algorithm for calculating the end-to-end bandwidth on a path and reserving required bandwidth so that a QoS route could be established. Instead of resorting to one single path to fulfill the needs of QoS, Liao et al. [10] proposed a multi-path QoS routing protocol, in which multiple paths are searched to support QoS together. In [19], Chen and Nahrstedt designed a distributed ticket-based QoS routing scheme that can deal with imprecise state information. However, the underlying MAC scheme is not specified and prioritization is not supported.In addition, no priority is supported. In [11], a framework of reliable routing is proposed by placing some reliable nodes in some special places in the network.

Location information has been leveraged to improve the performance of routing protocol [12–14, 24] and references therein). In [12, 13], the overhead associated with route discovery is greatly reduced since the search for a new route can be limited to a certain area. Basagni et al. [24] used location information to estimate the node mobility and take appropriate routing update frequency, leading to the decrease of update overhead. Further, location information can also be used to construct large and dense wireless ad hoc networks [14].

3. System model

We consider a mobile ad hoc network consisting of N nodes, among which there are M i-nodes. The system adopts TDMA as its channel access mechanism. As in [8], the transmission time scale is divided into frames, each of which contains a fixed number of time slots. Each frame is comprised of two phases. One is control phase and the other is data phase. In the control phase, all kinds of control functions are performed, such as pre-reservation request propagation, bandwidth reservation and connection setup, frame and slot synchronization. Each node may use a predefined slot to broadcast control messages to all of its neighbors. Data phase is used to transmit and receive data packets. For each time slot used to transmit or receive data (called data slot), it may be in one of the four states at one particular time: *FREE*, *PRE_RESERVED*, *USED_TX* and *USED_RX*. They indicate the state of the data slot being free, pre-reserved, used to transmit and used to receive, respectively. All connections considered in this paper are connection-oriented and requires constant bandwidth. In ad hoc networks using TDMA, bandwidth is measured in terms of time slots. Hereafter, the terms bandwidth and time slot are used interchangeably. A connection will specify its QoS requirement as the number of transmission time slots it needs prior to being admitted into the network. In the network, in addition to i-connections whose both endpoints are i-nodes, all the other connections are ordinary connections.

We assume in this work that each node is half-duplex, i.e., it cannot transmit and receive simultaneously. To successfully transmit a packet, both the transmitter and receiver nodes need to have one or more common time slots. Further restrictions apply to the selection of time slots along a path due to the hidden terminal problem. For example, in Fig. 2, if node n transmits packets destined to node 0 to node $n-1$ in time slot 1 and node $n-1$ forwards the packets to node $n-2$ in time slot 2. Since node $n-2$, two-hop away from node n, cannot hear from node n, it may also use time slot 1 to forward the packets to node $n-3$. Then, node $n-1$ will detect a collision and cannot correctly decode the packets from node n. We also assume that each node is aware of its own geographic location and time is synchronized at each node, as current Global Positioning System (GPS) [15, 16] can provide accurate location information and global timing. Finally, it is assumed that all i-nodes are aware of each other's geographic location. Since i-nodes need to frequently communicate with each other, they can piggyback their own location information and mobility information such as velocity and direction in data packets. I-nodes other than the source and the destination of i-connections can also acquire this information by overhearing. A more practical solution may be location registration and lookup service that maps node identities to location [17]. Further discussion on this issue is beyond the scope of this paper.

Formally, an ad hoc network is modeled as a graph $G = (N, L)$, where L is a set of bi-directional links [9]. A node i has a set of neighbors $NB_i = \{j \in N\colon (i, j) \in L\}$. Assume the set of data slots in one frame is $S = \{s_1, s_2, \ldots, s_K\}$. The set of slots TS_i in which node i transmits is defined as its transmission schedule. The set of nodes R_j^k, $R_j^k \in NB_i$, consists of the receivers in slot s_k, $s_k \in TS_i$. Set $RS_i = \{s_k \in S : i \in R_j^k, j \in NB_i\}$ is the set of slots in which node i uses to receive from its neighbors. Another two sets are defined for node i: $SRT_i = \{s_k \in S : s_k \notin TS_i, s_k \notin RS_i, \; s_k \notin \cup_{j \in NB_i} RS_j\}$, $SRR_i = \{s_k \in S : s_k \notin TS_i, \; s_k \notin RS_i, s_k \notin \cup_{j \in NB_i} TS_j\}$.$SRT_i$ is the set of slots in which i can transmit without causing interference to its current receiving neighbors and SRR_i is the set of slots in which i can receive without suffering interference from its current transmitting neighbors, given the current transmission schedule TS.

4. Bandwidth pre-reservation

To reduce the blocking probability of i-connections, we need to purposely pre-reserve some bandwidth beforehand. Before we proceed, it is important to note the differences in bandwidth reservation or pre-reservation in cellular networks and MANETs. In cellular networks, bandwidth is reserved at base stations; however, in ad hoc networks, bandwidth is reserved in each mobile node, which differs from base stations in several significant ways. First, a base station is a fixed infrastructure, which means once the bandwidth is reserved, it is always available to mobile users. Conversely, bandwidth pre-reserved in a mobile node may be unavailable as the mobile node moves. Second, a base station may have good knowledge about the mobility of all the mobile users in the cell it serves, and hence can make proper bandwidth reservation; in ad hoc networks, however, a mobile node does not know the mobility information about other mobile nodes, which makes it difficult to appropriately pre-reserve bandwidth. Finally, compared with a base station, a mobile node has limited power, processing capability, and buffer space; therefore, unlike a base station, a mobile node cannot afford any complicated algorithm for bandwidth pre-reservation.

4.1. Pre-reservation request propagation

Since statically pre-reserving bandwidth is not efficient due to node mobility or traffic dynamics, the bandwidth needs to be dynamically pre-reserved. Each node needs to know when to pre-reserve bandwidth, how much bandwidth to pre-reserve for future incoming i-connection requests and when to de-pre-reserve. To this end, each i-node of an i-node communication pair needs to provide the intermediate nodes with such information. For this purpose, a propagation and processing mechanism for pre-reservation requests is introduced.

As shown in Fig. 1, in order to pre-reserve some bandwidth along the path to i-node j, i-node i broadcasts a message called pre-reservation request to its neighbors, such as node n. The request contains the following information: (a) the

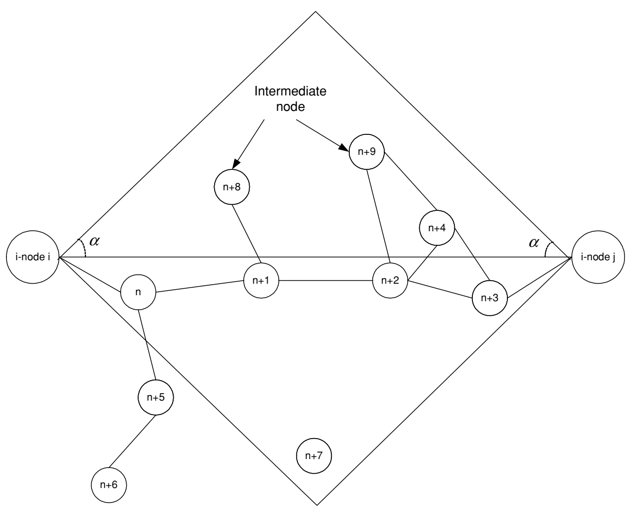


Fig. 1 Illustration of a quadrangle-shaped influence area

preferred number of time slots to be pre-reserved, *num_req* (it can be set by i-nodes depending on traffic history and application types); (b) the duration, *timer*, for which the time slots may be pre-reserved; (c) the location of both i-node *i* and i-node *j*; (d) i-node *i* and *j*'s identities (IDs); (e) a unique sequence number. Upon reception of the request, each one-hop neighbor of i-node *i*, like node *n*, will retrieve the location information of the source and the destination i-node. With this information, a quadrangle-shaped influence area will be built as shown in Fig. 1. The angle α adjusts the size of influence area and hence the control overhead incurred. If the neighbor happens to fall into this area, it will try to pre-reserve the number of time slots as required. For node *n*, this is the case. Assume the pre-reservation upper bound at node *n* is *UB*, and the current number of pre-reserved time slots is *num_tot*. Then, node *n* will take one of the following actions:

1) $num\_tot + num\_req \leq UB$, which means the node can pre-reserve *num_req* free time slots. Then the node will choose either *num_req* time slots or all currently free time slots denoted by *num_free*, whichever is smaller, and tag them with flag *PRE_RESERVED*. Then $num\_tot = num\_tot + min(num\_req, num\_free)$. The criterion governing how to choose which time slot to tag will be described in the following subsection. Besides, each chosen time slot will be associated with a timer, *slot_timer*(*k*), where *k* is the ID of the time slot. Initially, *slot_timer*(*k*) is set to 0. $slot\_timer(k) = timer$. When *slot_timer*(*k*) expires and the time slot will be freed, i.e., marked with flag *FREE* again.
2) $num\_tot = UB$, which means no more free time slots can be pre-reserved. Then the node will rank all the pre-reserved time slots in an increasing order of the value of their attached timer. The first *num_req* time slots' timer will be increased by a value of timer, i.e., $slot\_timer(k) = slot\_timer(k) + timer$, with the flag *PRE_RESERVED* unchanged. By doing this, though the number of pre-reserved time slots does not increase, their pre-reservation time is prolonged. Thus, this has the same effect as increasing the pre-reserved time slot number.
3) $num\_tot > UB$. Note that this may occur since we dynamically adjust *UB*. In this case, the node will choose the first *UB* time slots with the largest timer value. Among the *UB* time slots, *num_req* time slots with the least timer value will be increased by a value of timer. It can be seen that the $num\_tot - UB$ time slots that have small timer value may expire soon.
4) $num\_tot < UB$ and $num\_tot + num\_req > UB$. Then the node will choose $(UB - num\_tot)$ free time slots and pre-reserve them as described in 1). The left $(num\_req - (UB - num\_tot))$ time slots will be processed according to 2). Then $num\_tot = UB$.

Next, node *n* will broadcast the pre-reservation request message to its neighbors. In Fig. 1, they are node $n + 1$, $n + 5$. Then, node $n + 1$ and node $n + 5$ will also check their position. If they are also in this quadrangle, such as node $n + 1$, they will take the actions mentioned above. If they are not, such as node $n + 5$, they simply ignore the request. So node $n + 5$ will not pass on the pre-reservation request to node $n + 6$. To prevent a node from pre-reserving more than once for the same pre-reservation request, repeated reception of the same request will be detected by the IDs of source and destination i-nodes and the sequence number.

The frequency of sending out pre-reservation request is adjusted by a timer, $TIMER_{update}$, maintained at each i-node. Whenever this timer expires, the i-node will send another pre-reservation request. The adaptation of $TIMER_{update}$ will be addressed below. However, $TIMER_{update}$ is always a little smaller than *timer* to account for the propagation delay of pre-reservation requests. So the pre-reserved time slot at a node will not be freed just because the node cannot receive a new pre-reservation request to update the timer due to delay. Another implication is, once $TIMER_{update}$ is determined, *timer* is also determined. So far, it is assumed that the pre-reservation request is successfully transmitted in control phase. However, this mechanism still works even when a pre-reservation request may be lost or corrupted during transmission, since a new pre-reservation request will be transmitted some time later.

It is important to note that the pre-reserved time slots are not explicitly tied to the use of any two specific i-nodes, though they may be pre-reserved according to the requests by them. Accordingly, every i-connection can make use of them if they are available.

4.2. *Bandwidth pre-reservation criterion*

It is known that multi-hop topology of ad hoc networks allows spatial reuse of bandwidth. Different nodes can use the same bandwidth simultaneously as long as they are sufficiently separated. However, within one-hop and two-hop distance, time slots belonging to neighboring wireless links may collide with each other, due to the broadcasting nature of wireless transmission, half-duplex property, and the well-known hidden terminal problem. As a result, even though pre-reservation can set aside some free time slots at each node for i-connections, these slots, if carelessly pre-reserved, may be useless because of collisions. Worse yet, inappropriate pre-reservation will cause dramatic bandwidth underutilization since those improperly pre-reserved time slots may otherwise be available to other ordinary connections. We refer to Fig. 2 as an instance of pre-reservation failure. Source node *n* wants to open a connection with destination node 0. Prior to finding the path with sufficient bandwidth, suppose

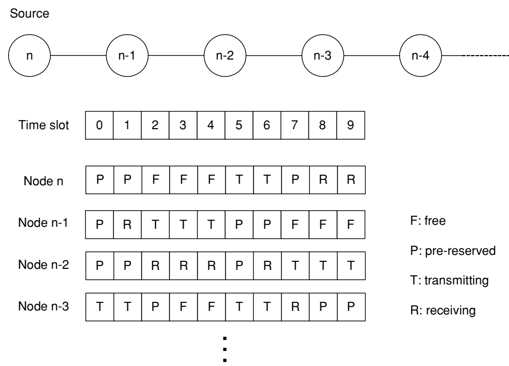


Fig. 2 Example of pre-reservation failure

the pre-reservation message is already sent from node n to node 0. As we see, node n pre-reserved slots 0, 1, 7; node $n-1$ pre-reserved slots 0, 5, 6; node $n-2$ pre-reserved slots 0, 1, 5; node $n-3$ pre-reserved slots 2, 8, 9. Although all these nodes have pre-reserved three time slots, due to time slot conflict, we cannot successfully find a path getting through from node n to node $n-3$ with one bandwidth unit (i.e., one slot), let alone a path from node n to node 0.

Moreover, as the connection path gets longer, it is less likely to find a path with sufficient bandwidth due to the transmission collision within one-hop and two-hop neighborhood. To illustrate this, we use a mathematical model to derive the relationship between the connection blocking probability and the hop length of the connection. We assume each connection requires the bandwidth of one time slot. For each link, the total bandwidth is S. The link bandwidth LB of each link is of i.i.d. Bernoulli distribution with success probability P_a, i.e., the probability mass function (PMF) of LB is defined as

$$B(S, i, p_a) = \binom{S}{i} p_a^i (1 - p_a)^{S-i} \tag{1}$$

where P_a indicates current traffic load in the network. We define $q(i) = B(S, i, p_a)$, for $i = 0, 1, \ldots, S$. We also define P_1 as the probability of $Pr(LB \geq 1)$, which is equal to $1 - B(S, 0, P_a)$. Based on this, we derive the connection blocking probability when the path comprises one, two and three hops, as shown in Appendix 1. We can see that, as the number of hops increases from 1 to 3, the blocking probability is also increased. For instance, if $S = 10$, $P_a = 0.1$, the blocking probabilities for QoS path with one, two, and three hops are 0.3487, 0.5908, 0.7695, respectively. If the number of hops of a path is greater than 3, the blocking probability becomes larger as time slot conflict increases.

Intuitively, if a wireless link can be isolated from the interference due to its neighboring links, more time slots at the link can be used due to the reduced collisions. Thus, more connections can be admitted into the network and hence a low blocking probability. Motivated by this intuition, we propose to pre-reserve time slots in such a way that, with the aid of each node's location information, the pre-reserved time slots are spread and hence they are not wasted simply because of the transmission collision within one-hop and two-hop neighborhood. Next, we detail the pre-reservation criterion in the following.

Assume the network geographically occupies a rectangular area with the size of $X \times Y\, m^2$. If the area is not a regular rectangle, it can be approximated with the smallest rectangle which can cover all the entire area. The transmission range of each node is Rm. A grid structure is built by dividing the network area into a grid of cells with size $2R \times 2R$. As a result, along the horizontal axis, there are $I = \lceil X/2R \rceil$ cells. And along the vertical axis, there are $J = \lceil Y/2R \rceil$ cells. The cells are denoted by (i, j), as shown in Fig. 3.

We further define the direction of a pre-reservation request. Assume when a node n, with location denoted by (x_n, y_n), receives a pre-reservation request originated by i-node s, with location denoted by (x_s, y_s), it will compare its horizontal distance from node s, i.e., $|x_n - x_s|$, with its vertical distance from node s, i.e., $|y_n - y_s|$. If $|x_n - x_s| \geq |y_n - y_s|$, we define this pre-reservation request to be horizontal for node n with respect to node s; otherwise, we define it to be vertical.

Next, time slots in a frame are also divided. Assume there are S data slots in one time frame. They are divided into two parts, denoted by H_1 and H_2, respectively. H_1 is the set of time slots which will be pre-reserved for horizontal pre-reservation requests and H_2 is the counterpart for vertical pre-reservation requests. To simplify notation, we also use H_1 and H_2 to refer to their respective size. The partition is in

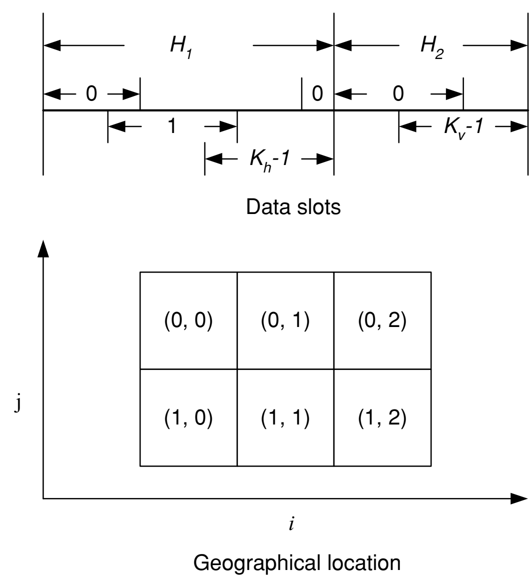


Fig. 3 Location and time division

proportion to the dimension of the geographic area, i.e.,

$$H_1 = \left\lfloor \frac{I}{I+J} S \right\rfloor \qquad H_2 = \left\lfloor \frac{J}{I+J} S \right\rfloor \tag{2}$$

As shown in Fig. 3, the time slots in H_1 are further equally divided into K_h ($K_h \geq 2$) sections, i.e., section $0, 1, 2, \ldots, K_h - 1$. Similarly, H_2 are further equally divided into K_v ($K_v \geq 2$) sections. The geometrical position of each cell is then mapped to the time axis. Along the horizontal direction, cell (i, j) is mapped to time slot section $mod(j, K_h)$. Along the vertical direction, cell (i, j) is mapped to time slot section $mod(i, K_v)$. Function $mod(p, q)$ returns the modulus obtained by dividing p into q. More specifically, the pre-reservation algorithm for node n, located in cell (i, j), to pre-reserve bw_num time slots for a pre-reservation request originated at i-node s, is outlined in Fig. 4, where N_s denotes the ID of the time slot section.

Under such a mapping rule, pre-reserved time slots are spread depending on the node's location and the direction of the pre-reservation request. Since the potential i-connection will likely follow the same shortest path experienced by pre-reservation requests, in most cases, the pre-reserved time slots for different i-connections will not collide with each other if they do not come in the same direction with respect to the node doing pre-reservation. For a connection with three or more hops, it is likely that any three of its consecutive hops span two neighboring cells and thus can have access to different time slots (By setting the cell dimension to $2R$, if there is a connection across the cell, there are at most two hops in the cell for the connection), as the pre-reserved time slots chosen from different sections will not overlap in most cases. Therefore, transmission collision among pre-reserved slots is reduced among different connections and different hops within one connection. As a result, when the QoS routing protocol is making bandwidth reservation along the path, it will find a path with sufficient bandwidth with a larger probability since the pre-reserved slots are available. Two important points are noted. First, there is some overlap, S_o, between every two adjacent sections. The reason why there is overlap in time slot section is to allow for the cases where some wireless links may span the boundary of two adjacent cells. In other words, one end of the link is located in one cell while the other end is located in an adjacent cell. With the overlapping, both node of the wireless link will be able to reserve the same time slot for transmission and reception. Second, as can be seen, this scheme is more effective when the network size is relatively large, e.g., for the network where QoS connections could traverse a few hops ($\geq$ 4 hops). Figure 5 shows the bandwidth pre-reservation following the above criterion, in which node n and $n-1$ belong to a cell, and node $n-2$ and $n-3$ belong to a neighboring cell. Similar to the case shown in Fig. 2, three time slots are pre-reserved. However, with our location-aware pre-reservation, it can be seen that the slots are pre-reserved such that we can find a path getting through from node n to node $n-3$, and to node 0 with one bandwidth unit.

Fig. 4 Pre-reservation criterion

Bandwidth Pre-reservation Algorithm

```
// assume node n needs to pre-reserve bw_num time slots for source i-node s
// location of s: L_s = (x_s, y_s); location of n in cell (i, j): L_n = (x_n, y_n);
while (bw_num)
{
    if (|x_n − x_s| ≥ |y_n − y_s|)        //pre-reserve time slots along horizontal direction
    {
    N_s = mod(j, K_h);
    if (⌊x_n/R⌋ == 0)        // in the left half of the cell
        randomly choose one free slot from the first half of N_s;
    else      // in the right half of the cell
        randomly choose one free slot from the second half of N_s;
    }
    else      // pre-reserve time slots along vertical direction
    {
    N_s = mod(i, K_v);
    if (⌊y_n/R⌋ == 0);       // in the upper half of the cell
        randomly choose one free slot from the first half of N_s;
    else      // in the lower half of the cell
        randomly choose one free slot from the second half of N_s;
    }
    bw_num–;
}
```

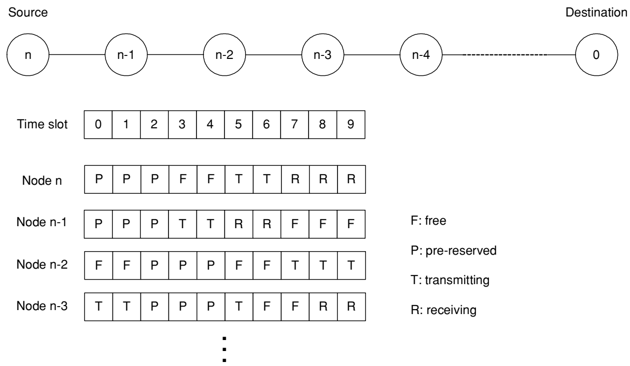


Fig. 5 An example of location-aware pre-reservation

4.3. *Upper bound of bandwidth to be pre-reserved*

An upper bound, *UB*, is used to ensure that bandwidth is not excessively pre-reserved, causing bandwidth under-utilization. To adapt to instantaneous network dynamics, such as traffic arrival information, node's geographic location, and node degree, it must be dynamically adjusted. Meanwhile, the upper bound provides the tradeoff between the performances of these two types of connections. If the upper bound is large, the i-connection blocking probability will be low while the blocking probability for ordinary connections will be high. Small upper bound implies just the opposite.

Each node in the network will monitor the incoming pre-reservation requests and the bandwidth requirement of each request. Let $TIMER_{avg}$ be the average of the timers associated with the pre-reserved time slots, and let λ_p be the arrival rate of the pre-reservation requests observed at a node. Besides, denote the average bandwidth requirement of each pre-reservation request by *BW*. Then, $TIMER_{avg} * \lambda_p * BW$ represents the average total bandwidth pre-reservation requirements for one node during the average pre-reservation period. Therefore, we set the upper bound to the following:

$$UB = TIMER_{avg}\lambda_p BW\Delta \tag{3}$$

where Δ $(0 < \Delta < 1)$ is a design parameter we introduce to adapt the upper bound *UB* so that the upper bound will not become too large. Thus, each node needs to estimate the current arrival rate of λ_p. Assume that each node measures the arrival rate at a fixed time period *T*, and the measured arrival rate at the *l*th $(l = 1, 2, \ldots)$ measurement period is $M_T(l)$. Then the arrival rate can be estimated using exponential moving average:

$$\lambda_p(l+1) = \beta\lambda_p(l) + (1-\beta)M_T(l) \tag{4}$$

where $M_T(l)$ can be obtained by

$$M_T(n) = \frac{\text{\# of new arrivals in } l\text{th period}}{T} \tag{5}$$

and β is a weighting factor, usually $0.5 < \beta < 1$. We can see that more weight is given to the arrival rates recently observed.

4.4. *Interval of sending pre-reservation requests*

The control overhead incurred by sending pre-reservation requests is determined by the sending interval, $TIMER_{update}$. At first glance, the overhead can be reduced by increasing $TIMER_{update}$. However, with large $TIMER_{update}$, bandwidth pre-reservation cannot quickly respond to the network topology change, which will cause performance degradation. Thus, $TIMER_{update}$ should be set so as to strike a balance between overhead and performance. Due to the fact that only i-nodes send out pre-reservation request, and they are not aware of intermediate nodes' mobility, our approach will adjust $TIMER_{update}$ based on the mobility of the sending i-node.

$$TIMER_{update} = MIN\left(\frac{R}{2V_{avg}}, TIMER_{max}\right) \tag{6}$$

where $TIMER_{max}$ is the maximum value that $TIMER_{update}$ can take on, V_{avg} is the average speed of the sending i-node. It can also be obtained using exponential moving average. $TIMER_{max}$ is introduced for two reasons. First, it can prevent $TIMER_{update}$ from being infinite, when the speed is very small. Second, even if the sending i-node is stationary, the intermediate nodes may move in or move out. So it ensures that the i-node sends pre-reservation requests to pre-reserve some time slots at the intermediate nodes.

4.5. *Control overhead analysis*

We assume that the network consists of *N* nodes uniformly distributed in an area of *A*. During an interval of $TIMER_{update}$, *num_data* packets are transmitted from i-node *i* to i-node *j*. If the average distance between two i-nodes is *D*, we can compute the area of the quadrangle as follows:

$$A_{quadrangle} = \frac{D^2}{2}\tan\alpha \tag{7}$$

If we count the control overhead as the number of nodes which sends or forwards the pre-reservation request, we obtain the control overhead per packet:

$$\begin{aligned} Overhead_{per\ packet} &= \frac{N * A_{quadrangle}}{A * num\_data} \\ &= \frac{ND^2\tan\alpha}{2A * num\_data} \end{aligned} \tag{8}$$

We can observe that the overhead per data packet increases with the increase of node density, the area of the quadrangle and decreases with increase of data packets sent during $TIMER_{update}$. This is verified by simulation in Section 6.

5. Location aware forwarding

5.1. *Bandwidth calculation and reservation*

To guarantee QoS, QoS routing protocols need to establish a path (or route) with sufficient bandwidth for admitting newly arriving connections. For any QoS connection originated at a node, the admission control test is as follows. The node first attempts to find a path by route discovery or check its own routing table, depending whether an on-demand routing protocol or a table driven routing protocol is used. If a path from the source to the destination is found, then the source will check if the path has sufficient bandwidth to accommodate the connection (for on-demand routing protocol, this is done with route discovery simultaneously). If yes, the bandwidth is reserved and the connection admitted; if no, the connection is rejected. The differential treatment of i-connection lies in the search for path bandwidth. For ordinary connections, they cannot reserve the time slots which have been pre-reserved for i-connections while i-connections can reserve either free or pre-reserved time slots. The bandwidth pre-reservation scheme described earlier ensures that each node's pre-reserved time slots are dispersed in time axis according to its location. Therefore, the chance that the discovered path has adequate bandwidth for i-connections is greatly increased as time slot collisions are reduced. i-connections thus will be admitted with higher priority than ordinary connections.

It can be seen from the above description that QoS routing protocol depends on the path bandwidth calculation algorithm to make admission decision. Unfortunately, in TDMA based link access schemes, the path bandwidth calculation problem is NP-complete [9]. We thus have to resort to heuristic approaches. *Forward Algorithm* (FA) is proposed as an efficient and distributed algorithm for calculating and reserving the end-to-end bandwidth along a path. Since every node is aware of its location, based on *FA*, we propose a *Location-Aware Forward Algorithm* (LAFA). The key idea is, when doing bandwidth reservation, LAFA will reserve the bandwidth with respect to the node's location according to the location-to-time-slot mapping criterion. In this way, the gain of reduced time slot collision due to the carefully designed location-dependent pre-reservation can be accrued. The algorithm is given in detail as follows.

Assume there is a path $P = \{n_m, n_{m-1}, \ldots, n_0\}$, where $n_i \in N$, $(n_i, n_{i-1}) \in L$, $i = m, m-1, \ldots, 1$. n_m is the source and n_0 is the destination. The set of time slots used on link (n_i, n_{i-1}) to support path $\{n_m, n_{m-1}, \ldots, n_k\}$ is defined as PB_i^k.

1) If $m = 1$,

$$PB_1^0 = LABW_1(node, IN, n) \tag{9}$$

2) If $m = 2$,

$$(PB_2^0, PB_1^0) = LABW_2(LB_2, LB_1); \tag{10}$$

3) If $m \geq 3$,

$$(PB_m^{m-2}, PB_{m-1}^{m-2}) = LABW_2(LB_m, LB_{m-1}); \tag{11}$$

for $k = m - 3$ to 0 do

$$(PB_{k+3}^k, PB_{k+2}^k, PB_{k+1}^k) = LABW_3(PB_{k+3}^{k+1}, PB_{k+2}^{k+1}, LB_{k+1}); \tag{12}$$

where $LB_i = SRT_i \cap SRR_{i-1}$, denoting the link bandwidth of link (n_i, n_{i-1}).

The end-to-end bandwidth of path P is $|PB_1{}^0|$. Note in the above algorithm, functions $LABW_1$, $LABW_2$, and $LABW_3$ are given in Appendix 2.

Since each node has pre-reservations, we need to modify the two sets, namely, SRT_i and SRR_i, which are used in FA. For i-connections, SRT_i is defined as $\{s_k \in S : s_k \notin TS_i, s_k \notin RS_i, s_k \notin \cup_{j \in NB_i} RS_j\}$ and SRR_i is defined as $\{s_k \in S : s_k \notin TS_i, s_k \notin RS_i, s_k \notin \cup_{j \in NB_i} TS_j\}$ for ordinary connections, SRT_i is defined as $\{s_k \in S : s_k \notin TS_i, s_k \notin RS_i, s_k \notin \cup_{j \in NB_i} RS_j, s_k \notin \cup_{j \in NB_i} PRV_j\}$ and SRR_i is defined as $\{s_k \in S : s_k \notin TS_i, s_k \notin RS_i, s_k \notin \cup_{j \in NB_i} TS_j, s_k \notin \cup_{j \in NB_i} PRV_j\}$, where PRV_i, the pre-reservation set at node i, is defined as $\{s_k \in S : s_k \text{ is tagged } PRE\_RESERVED\}$.

5.2. *Impact on re-routing*

In QoS routing algorithms, an alternate path needs to be searched and used when the primary path is broken due to mobility. Since each i-node pair has asked the intermediate node in the influence area to pre-reserve bandwidth, this quadrangle may contain several alternate paths. Therefore, when the primary path is not available, the routing algorithm can easily find an alternate path with sufficient bandwidth. Therefore, the re-routing successful probability will be increased as well as the connection acceptance ratio, which will be shown in Section 6.

6. Performance evaluation

In this section, the performance of our proposed scheme is evaluated through extensive simulations. Our simulation study is carried out using OPNET Modeler 8.0. An ad hoc network consisting of 80 mobile nodes is simulated in a 1600 $\times$ 400 m^2 area. The transmission range of a node is 200 m, so a long QoS connection may traverse more than 4 hops. The initial position of each mobile node is uniformly distributed in the entire network. Their mobility follows the waypoint mobility model [23]. That is, after remaining stationary for a period of pause time, a node randomly chooses a destination and starts to move toward it. When moving, a node will randomly choose a speed from 1–10 m/s. It can be seen that the shorter the pause time, the higher the mobility of each node. There are 8 i-nodes, i.e., 10% of all nodes. Connection requests arrive following Poisson process, each requiring one data slot in each frame. The duration of each connection is exponentially distributed with mean of 60 sec. In each time frame, the data slot is 5 ms and the control slot is 0.1 ms. We assume there are 30 control slots in control phase and 72 data slots in data phase. So the frame length is 30*0.1 + 72*5 = 363 ms. We also assume a data packet can be transmitted in one data slot. The angle α is 45 degree unless otherwise stated. The simulation duration is 900 sec.

For comparison purpose, we consider three schemes in our simulation. The first one is the scheme without any pre-reservation. The second is our proposed scheme, which makes bandwidth pre-reservation in a quadrangle area for i-connections. We also created the last one that adopts bandwidth pre-reservation. However, it only pre-reserves time slots along the shortest path between any two i-nodes. Note that depending on which routing protocol is used, the shortest path information may or may not be available when an i-node sends out the pre-reservation message. In the simulation, we assume that this scheme knows this information. In the figures presented below, these three are designated as *No pre-reservation*, *Quadrangle* and *SP*, respectively.

Through simulation results, we observe that, when the network traffic load is light, the performances of all the three schemes are very close. This is because that, under light traffic conditions, the network always has enough bandwidth to accommodate new connections. Therefore, the connection blocking probabilities are all very low. However, it is even more critical for these schemes to perform well under medium to heavy traffic load. For this reason, we mainly focus on these schemes' performance in a relatively heavy traffic situation, where the average connection arrival rate at each node is 0.2 connection/sec. All results presented here are averaged over 10 simulation runs.

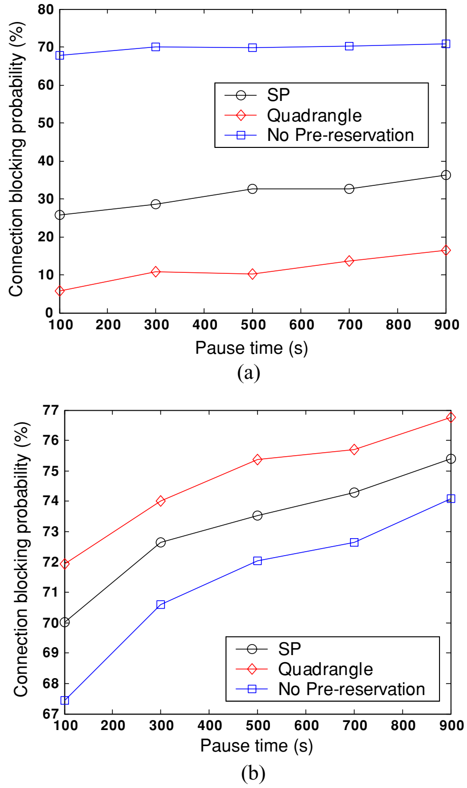


Fig. 6 Connection blocking probability (a) i-connection, (b) ordinary connection

Figure 6 shows the performances of the three schemes in terms of connection blocking probability. For i-connection, our scheme provides the lowest blocking probability among all three schemes. This is expected as each i-node is sending pre-reservation requests to pre-reserve some time slots at the intermediate nodes within the influence area. Scheme *SP* is worse since it only pre-reserves slots at nodes on the shortest path, which may be unavailable due to mobility. We also observe that as the mobility is increased, the connection blocking probability is slightly reduced. This is not unexpected due to the following fact. When mobility increases, the ongoing connection may be dropped, releasing the bandwidth occupied for newly incoming connections. Therefore, the connection blocking probability drops instead of increasing. For ordinary connections, the performance of scheme *No pre-reservation* is the best. This is reasonable since the entire network bandwidth is fixed, as more bandwidth is given to i-connection, less bandwidth is left for accommodating ordinary connections. However, the differences are very small.

Next, we investigate the performances of these three schemes when re-routing is considered. Figure 7 presents the re-routing successful probabilities for both types of

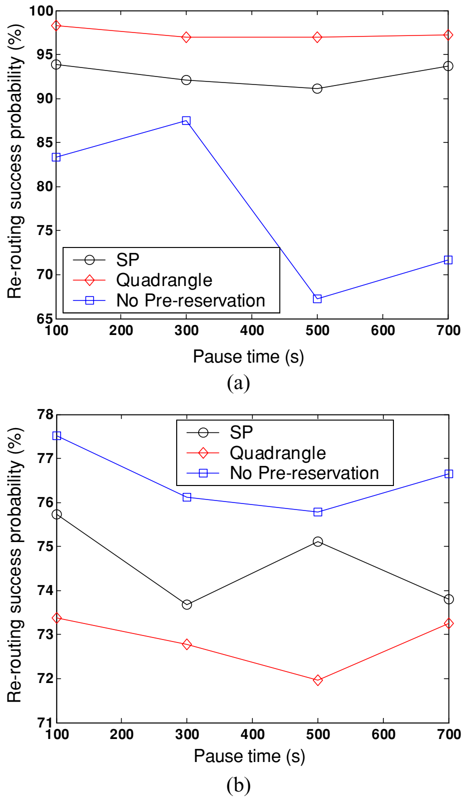


Fig. 7 Re-routing success probability (a) i-connection, (b) ordinary connection

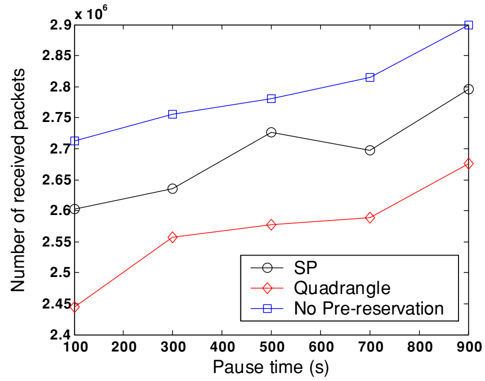


Fig. 8 Throughput

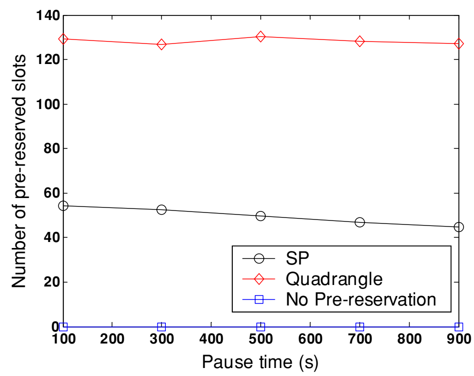


Fig. 9 Average number of pre-reserved time slots

connections. Again, for i-connections, our scheme, *Quadrangle* shows the best performance, since it pre-reserves bandwidth in the quadrangle. When one connection is dropped due to mobility, it will have high probability to find another path with sufficient bandwidth in the quadrangle. *SP* is worse and *No pre-reservation* is the worst. For ordinary connections, their performances are in an opposite order for the same reason described earlier.

Figure 8 illustrates the throughput performance in terms of successfully received packets. For all three schemes, throughput increases as mobility decreases. This shows the negative effect of mobility on the throughput: higher mobility causes more ongoing connections to be dropped and in turn reduce the throughput. On the other hand, we observe that throughput is close to one another among the three schemes, which means, although our scheme adopts bandwidth pre-reservation for i-connection, the network throughput is not seriously affected.

The number of pre-reserved time slots in the network is presented in Fig. 9. As expected, our scheme pre-reserves more time slots compared with the other two. However, considering the entire network has a number of time slot 72*80 = 5760, we only need to pre-reserve about 2.4%. Figure 9 also illustrates a desirable feature of the scheme: the number of pre-reserved time slots is insensitive to the changes in mobility, which ensures the scheme does not pre-reserve too many time slots in high mobility scenarios.

Figure 10 shows the control overhead incurred by sending pre-reservation requests. We count the overhead as the number of times for which each node, either i-nodes or ordinary nodes, broadcasts the pre-reservation request. We see that the overhead is small and decreases with mobility, showing the adaptability of our scheme to mobility. With the lowest blocking probability and highest re-routing successful probability for i-connections and the small control overhead incurred, our scheme, i.e.,scheme *Quadrangle* is able to provide high priority communication service to i-nodes with QoS guarantee at a low cost. Scheme *SP* seems to perform fairly well, however, it cannot be used with some on-demand routing protocols as the shortest path information is not available for bandwidth pre-reservation.

Finally, we study the impact of angle α on the performance of scheme *Quadrangle*, as shown in Figs. 11–15. In Figs. 11 and 12, for i-connection, as α increases, the connection blocking probability is decreased and re-routing successful probability is increased. This is expected, since larger α implies larger quadrangle, which in turn will have more time slots pre-reserved for i-connections. For ordinary connections, large α has an opposite effect. It can also be

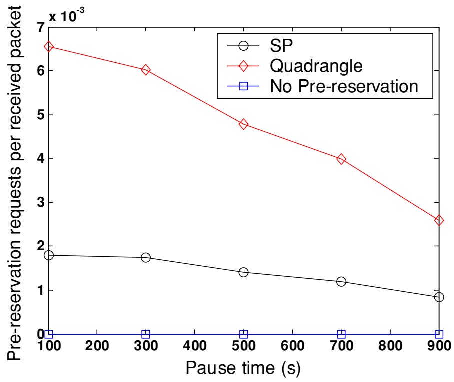


Fig. 10 Pre-reservation requests per received packet

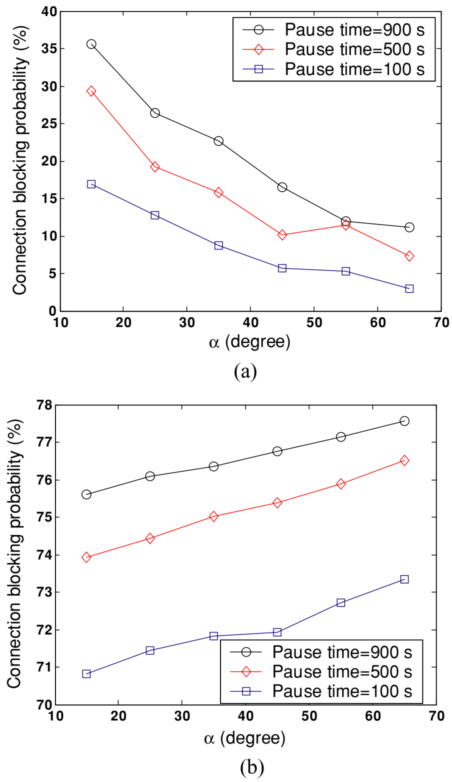


Fig. 11 Connection blocking probability (a) i-connection, (b) ordinary connection

seen that the superiority of our proposed scheme is still maintained.

Figure 13 shows the throughput is slightly affected by the increase in α. Also, more overhead is incurred to propagate the pre-reservation requests and more time slots need to be pre-reserved for i-connections, as α increases, as shown in Figs. 14 and 15. Therefore, α can be used as a design parameter determining the extent to which i-connection should be granted high priority. Figure 15 also verifies the insensitivity of the scheme to the changes in mobility, which is highly desirable as over-pre-reservation in the case of high mobility is avoided.

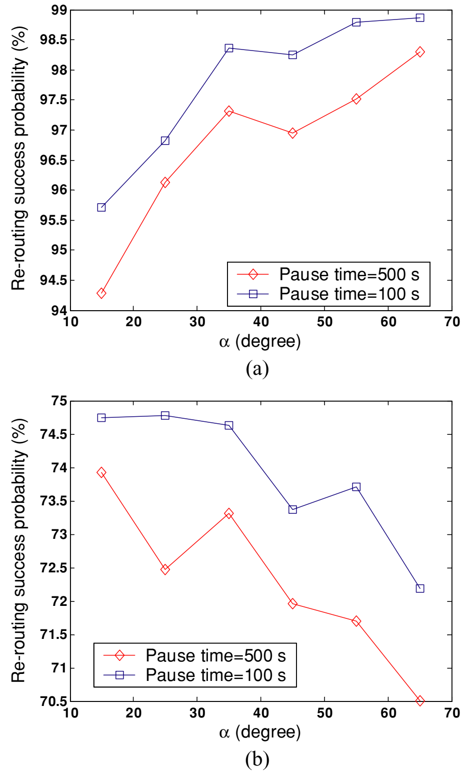


Fig. 12 Re-routing success probability (a) i-connection, (b) ordinary connection

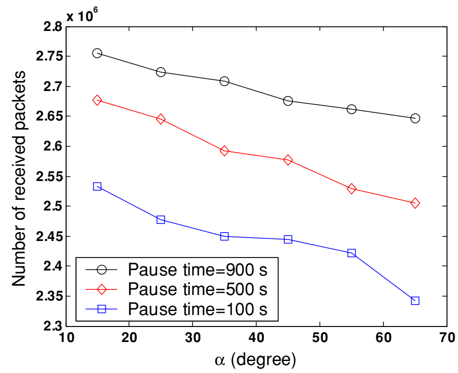


Fig. 13 Throughput

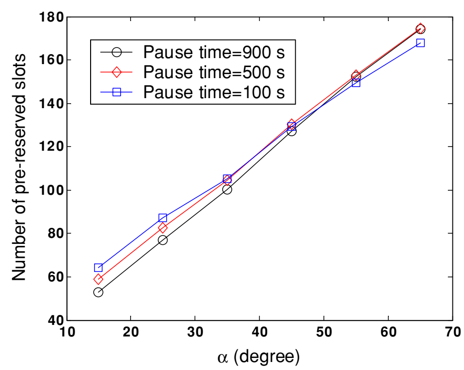


Fig. 14 Average number of pre-reserved time slots

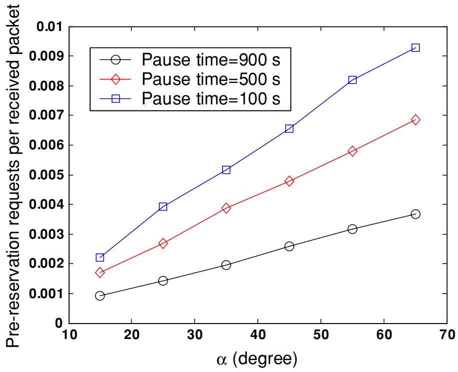


Fig. 15 Pre-reservation requests per received packet

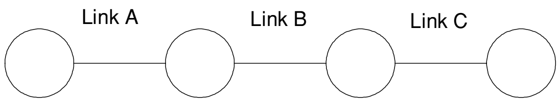


Fig. 16 A path for derivation

7. Discussions and future work

In this section, we comment on several design issues and discuss future work.

7.1. *Mobility*

Through the way that a quadrangle-shaped influence area is formed between any two i-nodes, an i-node pair can always pre-reserve some time slots in the intermediate nodes in between it. For intermediate nodes, when they are located in the influence area, they will always pre-reserve some time slots for possible communication between the two i-nodes; when they move out, the pre-reserved time slots for the i-node pair will be freed, since future pre-reservation requests from the two i-nodes may not be received for updating the timers associated with the time slots. When one of the two i-nodes moves, so will the quadrangle, which will be generated according to the new location of the two i-nodes. Therefore, this mechanism suits well with the mobile nature of ad hoc networks. Of course, QoS is only feasible in the network that is static or with low-to-medium mobility; when the network topology changes too fast, there is no way to support QoS [9,19].

7.2. *Trade off between overhead and performance*

The control overhead in bandwidth pre-reservation is mainly the propagation of pre-reservation requests, which consumes network bandwidth. Since there are M i-nodes in the network, they may form up to $M(M-1)$ half-duplex communication pairs, each of which sends pre-reserving request according its own sending interval. Note that we need to distinguish the direction of communication for each i-node pair, since each direction will demand a QoS connection to be set up before sending data. Thus, the overhead depends on the number of i-nodes in network, the geographical location of each i-node, as well as the node density and the traffic load between i-nodes during $TIMER_{update}$, which are shown in the overhead analysis in Section 4. It is important to note that when two i-nodes are far away from each other and have a large amount of data to transmit, the quadrangle bandwidth pre-reservation may involve many nodes if the node density is high. Consequently, the control overhead may be high. In this case, we may adopt the strategy of *hop limit* [25]. That is, each pre-reservation request contains a field of "hop limit" Every time the request is forwarded by a node, this limit is decremented; if it becomes zero, the request is simply discarded. In this way, we limit the number of intermediate nodes allowed to forward the request and hence reduce the overhead. Moreover, we can also adapt the angle α to the distance between two i-nodes and the node density in the network.

Nevertheless, the overhead is relatively small for the following reasons. First, it is anticipated only a very small percent of nodes are i-nodes in ad hoc or sensor networks. Second, as we can see from the estimation of $TIMER_{update}$, it adapts to node mobility. If nodes are static or move at a low to medium speed, the overhead can be further reduced as $TIMER_{update}$ will be increased. It is worth noting when estimating $TIMER_{update}$, we have not considered the traffic type or traffic history of each i-node. In reality, for those i-nodes that transmit data less frequently, $TIMER_{update}$ can be further adapted to reduce overhead. We intend to further explore this aspect in the future.

8. Conclusion

Since neither routing nor MAC protocols can effectively provide high priority communications with QoS guarantee for some important nodes in mobile ad hoc networks, we proposed a novel cross-layer approach for mobile ad hoc networks in this paper. We first propose a location-aware bandwidth pre-reservation mechanism to pre-reserve bandwidth for connections between important nodes by utilizing each node's geographic location information, thereby reducing the potential transmission collisions. Then, a location-aware forward algorithm (LAFA) is proposed to calculate and reserve end-to-end bandwidth for such high priority connections. In this way, our scheme can not only reduce the transmission collision and hence increasing resource utilization, but also adapt to network topology changes due to mobility. Therefore, high priority communication services between important nodes can be provided by the network with high probability and QoS guarantees without incurring too much overhead and severely blocking other connections, a goal that cannot be achieved without the collaboration between the routing layer and the MAC layer and the use of location information. Finally, extensive simulation verifies the performance of our proposed scheme.

Acknowledgments

This work was supported in part by the US Office of Naval Research under grant N000140210464 (Young Investigator Award) and under grant N000140210554, and by the US National Science Foundation under grant ANI-0093241 (CAREER Award) and under grant ANI-0220287.

Appendix 1

Derivation of connection blocking probability

1) For a path with only one hop, it is obvious that the connection blocking probability PB_1 is equal to the probability that $LB = 0$, i.e., $PB_1 = q(0)$.
2) For a path with two hops, say link A and B, first we consider the connection acceptance probability PA_2. We know that a connection request will be accepted if both links have $LB \geq 1$, the probability of which is P_1^2. However, in this case, there is one scenario that we cannot build the connection, which is equal to $q^2(1)/S$. So the probability of acceptance is $P_1^2 - q_2(1)/S$. Therefore, the blocking probability is $PB_2 = 1 - (P_1^2 - \frac{q^2(1)}{S})$.
3) For a path with three hops, we also first calculate the connection acceptance probability PA_3. The probability that all three hops have at least one time slot available is P_1^3. However, we need to subtract from this probability several possibilities.
   a) Any two have the same time slot only. The probability is $\frac{3q^2(1)P_1S-2q^3(1)}{S^2}$.
   b) One link has two time slots and the other two links each have one of the two time slots. The probability is $\frac{6q(2)q^2(1)}{S^2}$.
   c) Two links have the same two time slots and the rest one link has one of the two time slots. The probability is $\frac{12q^2(2)q(1)}{S^2(S-1)}$.
   d) All the three links have the same two time slots. The probability is $\frac{4q^3(2)}{S^2(S-1)^2}$.

It can be shown that there are no other cases where a connection cannot be admitted. Therefore, PA_3 is the probability by subtracting all these probabilities from P_1^3. Therefore, the blocking probability is $PB_3 = 1 - (P_1^3 - \frac{3q^2(1)P_1S-2q^3(1)+6q(2)q^2(1)}{S^2} - \frac{12q^2(2)q(1)}{S^2(S-1)} - \frac{4q^3(2)}{S^2(S-1)^2})$.

Appendix 2

Functions used in LAFA

```
function(OUT) = BW_location(node)
  get node's location (x_n, y_n) in cell (i, j);
  get source's location (x_s, y_s);
  if |x_n − x_s| ≥ |y_n − y_s|
    N_time section = mod(j, K_h);
    OUT = half of all the pre-reserved or free time slots in
        time section N_time section;
  else
    N_time section = mod(i, K_v);
    OUT = half of all the pre-reserved or free time slots in
        time section N_time section;
    return;
  function(OUT) = LABW_1(node, IN, n)
    assert(n ≤ |IN|);
    S = BW_location(node);
    E_1 = S ∩ IN;
    E_2 = S̄ ∩ IN;
    if  |E_1| ≥ n
    randomly choose n elements from IN as OUT;
    return;
  else
    randomly choose n − |E_1| elements from E_2 as E_3;
    OUT = E_1 ∩ E_3;
    return;
  function(OUT_2, OUT_1) =
        LABW_2(node_2, node_1, IN_2, IN_1)
  C = IN_1 ∩ IN_2;
  E_1 = IN_1 ∩ ‾IN_2‾;
  E_2 = IN_2 ∩ ‾IN_1‾;
  if |E_2| ≥ |IN_1|
    OUT_1 = IN_1 ;
    OUT_2 = LABW_1(node_2, E_2, |IN_1|);
    return;
  elseif |E_1| ≥ |IN_2|
    OUT_1 = LABW_1(node_1, E_1, |IN_2|);
    OUT_2 = IN_2;
    return;
  else
    T = floor(|IN_1 ∪ IN_2|/2);
    C_2 = LABW_1(node_2, C, T − |E_2|);
    C_1 = C ∩ ‾C_2‾;
    OUT_1 = LABW_1(node_1, C_1 ∪ E_1, T);
    return;
  function(OUT_3, OUT_2, OUT_1 ) =
        LABW_3(node_3, node_2, node_1, IN_3, IN_2, IN_1)
    assert(|IN_3| = |IN_2| &&IN_2 ∩ IN_3 = φ);
    C_21 = IN_2 ∩ IN_1;
    C_31 = IN_3 ∩ IN_1;
```

$E_1 = IN_1 \cap \overline{C_{21}} \cap \overline{C_{31}};$
$E_2 = IN_2 \cap \overline{C_{21}};$
$E_3 = IN_3 \cap \overline{C_{31}};$
$if\ \ |E_1| \geq |IN_2|$
$OUT_1 = LABW_1(node_1, E_1, |IN_2|);$
$OUT_2 = IN_2;$
$OUT_3 = IN_3\ ;$
return;
$elseif\ |E_3| \geq |LABW_2(node_2, node_1, IN_2, IN_1)|$
$(OUT_2, OUT_1\ ) = LABW_2\ (node_2, node_1, IN_2, IN_1\ );$
$OUT_3 = LABW_1\ (node_3, E_3, |OUT_1|);$
return;
$elseif\ |E_2| \geq |LABW_2(node_3, node_1, IN_3, IN_1)|$
$(OUT_3, OUT_1\ ) = LABW_2\ (node_3, node_1, IN_3, IN_1\ );$
$OUT_2 = LABW_1\ (node_2, E_2, |OUT_1|);$
return;
else
$T = floor(|IN_1 \cup IN_2 \cup IN_3| / 3);$
$C_{31}^3 = LABW_1(node_3, C_{31}, T - |E_3|);$
$C_{31}^1 = C_{31} \cap \overline{C_{31}^3};$
$C_{21}^2 = LABW_1(node_2, C_{21}, T - |E_2|);$
$C_{21}^1 = C_{21} \cap \overline{C_{21}^2};$
$OUT_1 = LABW_1(node_1, E_1 \cup C_{21}^1 \cup C_{31}^1, T);$
$OUT_2 = E_2 \cup C_{21}^2;$
$OUT_3 = E_3 \cup C_{31}^3;$
return;

References

[1] A. Veres, A.T. Campbell, M. Barry and L.H. Sun, Supporting service differentiation in wireless packet networks using distributed control, IEEE Journal on Selected Areas in Communications (Oct. 2001) 2094–2104.

[2] G.S. Ahn, A.T. Campbell, A. Veres and L.H. Sun, SWAN: service differentiation in stateless wireless ad hoc networks, in: *Proceedings of IEEE INFOCOM* (June 2002).

[3] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi and E. Knightly, Distributed Multi-hop scheduling and medium access with delay and throughput constraints, in: *Proceedings of ACM MOBICOM* (July 2001).

[4] I. Chlamtac, A. Farago and H. Zhang, Time-spread multiple-access (TSMA) protocols for multihop mobile radio networks, IEEE/ACM Transactions on Networking (Oct. 1997) 804–812.

[5] A.M. Chou and V.O.K. Li, Slot allocation strategies for TDMA protocols in multihop packet radio networks, in: *Proceedings of IEEE INFOCOM* (May 1992).

[6] P. Bjorklund, P. Varbrand and D. Yuan, Resource optimization of spatial TDMA in ad hoc radio networks: a column generation approach, in: *Proceedings of IEEE INFOCOM* (March–April 2003).

[7] C.R. Lin and J.-S. Liu, QoS routing in ad hoc wireless networks, IEEE Journal on Selected Areas in Communications (Aug. 1999) 1426–1438.

[8] C.R. Lin, Admission control in time-slotted mulithop mobile networks, IEEE Journal on Selected Areas in Communications (Oct. 2001) 1974–1983.

[9] C. Zhu and M.S. Corson, QoS routing for mobile ad hoc networks, in: *Proceedings of IEEE INFOCOM* (June 2002).

[10] W.H. Liao, Y.C. Tseng, S.L. Wang and J.P. Sheu, A multi-path QoS routing protocol in a wireless mobile ad hoc network, in: *Proceedings of IEEE ICN* (July 2001).

[11] Z. Ye, S.V. Krishnamurthy and S.K. Tripathi, A framework for reliable routing in mobile ad hoc networks, in: *Proceedings of IEEE INFOCOM* (March–April 2003).

[12] Y. Ko and N. Vaidya, Location-aided routing (LAR) in mobile ad hoc networks, in: *Proceedings of ACM MobiCom* (Oct. 1998).

[13] R. Banka and G. Xue, Angle routing protocol: location aided routing for mobile ad-hoc networks using dynamic angle selection, in: *Proceedings of IEEE MILCOM* (Oct. 2002).

[14] B. Karp and H. Kung, GPSR: Greedy perimeter stateless routing for wireless networks, in: *Proceedings of ACM MobiCom* (Aug. 2000).

[15] USCG Navigation Center GPS page, http://www.navcen.uscg.gov/gps/default.htm, May 2003.

[16] B.W. Parkinson and S.W. Gilbert, NAVSTAR: global positioning system—ten years later, Proceedings of the IEEE (1983) 1177–1186.

[17] J. Li, J. Jannotti, D. DeCouto, D. Karger and R. Morris, A scalable location service for geographic ad hoc routing, in: *Proceedings of ACM MobiCom* (Aug. 2000).

[18] D. Hong and S.S. Rappaport, Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures, IEEE Transactions on Vehicular Technology (Aug. 1986) 77–92.

[19] S. Chen and K. Nahrstedt, Distributed quality-of-service routing in ad hoc networks, IEEE Journal on Selected Areas in Communications (Aug. 1999) 1488–1505.

[20] S. Nesargi and R. Prakash, Distributed wireless channel allocation in networks with mobile base stations, IEEE Transactions on Vehicular Technology (Nov. 2002) 1407-1421.

[21] W.H. Liao, Y.C. Tseng and K.P. Shih, A TDMA-based bandwidth reservation protocol for QoS routing in a wireless mobile ad hoc network, in: *Proceedings of IEEE ICC* (April–May 2002).

[22] I. Ada and C. Castelluccia, Differentiation mechanisms for IEEE 802.11, in: *Proceedings of IEEE INFOCOM* (April 2001).

[23] J. Yoon, M. Liu and B. Noble, Sound mobility models, in: *Proceedings of ACM MobiCom* (Sept. 2003).

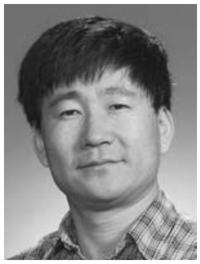
[24] S. Basagni, I. Chlamtac, V. Syrotiuk and B. Woodward, A distance routing effect algorithm for mobility (DREAM), in: *Proceedings of ACM MobiCom* (Oct. 1998).

[25] D. Johnson, D. Maltz and J. Broch, DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks, in Ad Hoc Networking, edited by Charles E. Perkins, Chapter 5, pp. 139–172, Addison-Wesley 2001.

Xiang Chen received the B.E. and M.E. degrees in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1997 and 2000, respectively. Afterwards, he worked as a MTS (member of technical staff) in Bell Laboratories, Beijing, China. He is currently working toward the Ph.D. degree in the department of Electrical and Computer Engineering at the University of Florida. His research is focused on protocol design and performance evaluation in wireless networks, including cellular networks, wireless LANs, and mobile ad hoc networks. He is a member of Tau Beta Pi and a student member of IEEE.

Wei Liu received the BE and ME degrees in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 1998 and 2001, respectively. He is currently pursuing the P.hD. degree in the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where he is a research assistant in the Wireless Networks Laboratory (WINET). His research interest includes QoS, secure and power efficient routing, and MAC protocols in mobile ad hoc networks and sensor networks. He is a student member of the IEEE.

Hongqiang Zhai received the B.E. and M.E. degrees in electrical engineering from Tsinghua University, Beijing, China, in July 1999 and January 2002 respectively. He worked as a research intern in Bell Labs Research China from June 2001 to December 2001, and in Microsoft Research Asia from January 2002 to July 2002. Currently he is pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Florida. He is a student member of IEEE.



Yuguang Fang received a Ph.D. degree in Systems and Control Engineering from Case Western Reserve University in January 1994, and a Ph.D. degree in Electrical Engineering from Boston University in May 1997. From June 1997 to July 1998, he was a Visiting Assistant Professor in Department of Electrical Engineering at the University of Texas at Dallas. From July 1998 to May 2000, he was an Assistant Professor in the Department of Electrical and Computer Engineering at New Jersey Institute of Technology. In May 2000, he joined the Department of Electrical and Computer Engineering at University of Florida where he got the early promotion to Associate Professor with tenure in August 2003 and to Full Professor in August 2005. He has published over 180 papers in refereed professional journals and conferences. He received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He is currently serving as an Editor for many journals including IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEETransactions on Mobile Computing, and ACM Wireless Networks. He is also actively participating in conference organization such as the Program Vice-Chair for IEEE INFOCOM'2005, Program Co-Chair for the Global Internet and Next Generation Networks Symposium in IEEE Globecom'2004 and the Program Vice Chair for 2000 IEEE Wireless Communications and Networking Conference (WCNC'2000).