DELAR: A Device-Energy-Load Aware Relaying Framework for Heterogeneous Mobile Ad Hoc Networks

Wei Liu, *Member, IEEE*, Chi Zhang, *Student Member, IEEE*, Guoliang Yao, *Member, IEEE*, and Yuguang Fang, *Fellow, IEEE*

***Abstract*—This paper addresses energy conservation, a fundamental issue of paramount importance in heterogeneous mobile ad hoc networks (MANETs) consisting of powerful nodes (i.e., P-nodes) as well as normal nodes (i.e., B-nodes). By utilizing the inherent device heterogeneity, we propose a cross-layer designed Device-Energy-Load Aware Relaying framework, named *DELAR*, to achieve energy conservation from multiple facets, including power-aware routing, transmission scheduling and power control. In particular, we design a novel power-aware routing protocol that nicely incorporates device heterogeneity, nodal residual energy information and nodal load status to save energy. In addition, we develop a hybrid transmission scheduling scheme, which is a combination of reservation-based and contention-based medium access control schemes, to coordinate the transmissions. Moreover, the novel notion of "mini-routing" is introduced into the data link layer and an Asymmetric MAC (A-MAC) scheme is proposed to support the MAC-layer acknowledgements over unidirectional links caused by asymmetric transmission power levels between powerful nodes and normal nodes. Furthermore, we present a multi-packet transmission scheme to improve the end-to-end delay performance. Extensive simulations show that DELAR can indeed achieve energy saving while striking a good balance between energy efficiency and other network performance metrics.**

***Index Terms*—Energy Conservation; Routing; MAC; Power control; Heterogeneity; MANETs.**

I. INTRODUCTION

A MANET usually consists of battery-powered mobile devices (nodes) which will become useless once their limited power is depleted, leading to network performance degradation, network partition, or potential network collapse. As a result, the energy conservation of mobile devices becomes a crucial issue for normal operations of MANETs.

There are intensive research on energy efficient protocols in the literature [1]. Most research works focus on the MANETs in which all nodes are treated identical. However, the heterogeneity of mobile devices seems to be inherent and has been commonly observed in MANETs [2]. For instance, mobile devices in the same network may differ in their CPU speed, available memory, operating systems, or protocol stacks. They may also have various power resources or transmission capability[1], communication capacities, traffic patterns, or mobility patterns [3]. Such heterogeneity, coupled with error-prone and time-varying wireless channels and dynamic changing network topologies, further complicates the issue on energy conservation in ad hoc networks. For example, different transmission power levels between two communication nodes may result in unidirectional links, which would restrict the direct application of many network protocols such as IEEE 802.11, which assume bidirectional and symmetric links. However, just as everything always has two sides, such heterogeneity also introduces opportunities for developing more efficient energy conservation techniques. How to design energy efficient network protocols by taking advantage of these heterogeneities is important and challenging, which is the main focus of this paper.

In this paper, we focus on the heterogeneous ad hoc networks, where most nodes, denoted as *B-nodes*, are equipped with limited power sources like batteries, while some other nodes, denoted as *P-nodes*, have relatively unlimited power supplies, e.g., power scavenging units such as solar cells, or dynamos when they are installed in mobile vehicles, etc. The basic idea is to develop more energy conscious protocols by taking advantage of the heterogeneity of mobile devices, i.e., being generous in using the P-nodes while conservative in using the B-nodes. More specifically, our contributions are mainly fourfold. First, following the cross-layer design philosophy, we propose a Device-Energy-Load Aware Relaying framework, named *DELAR*, to achieve energy conservation by utilizing the inherent heterogeneity of nodal power capabilities. Second, we design a hybrid transmission scheduling scheme, combining both the reservation-based and contention-based medium access control schemes, to coordinate the transmissions among P-nodes and B-nodes, which attempts to make the best use of powerful nodes while controlling their interferences to other ongoing transmissions. Third, we develop "mini-routing" and asymmetric MAC (*A-MAC*) protocols to support the MAC layer acknowledgements over unidirectional links due to the use of asymmetric transmission power levels between P-nodes and B-nodes. Finally, we present a multi-

Manuscript received 1 October 2010; revised 15 February 2011. This work was partially supported by NSF under grants CNS-0916391 and CNS-0716450. The work of Y. Fang was also partially supported by the National Natural Science Foundation of China under grant No. 61003300.

W. Liu is with Scalable Network Technologies, Inc., Los Angeles, CA.

C. Zhang, G. Yao and Y. Fang are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA. (e-mail: {zhangchi@,guoliang@,fang@ece.}ufl.edu). Y. Fang was a Changjiang Scholar Chair Professor with the National Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China.

Digital Object Identifier 10.1109/JSAC.2011.110907.

[1]In this paper, whenever appropriate, we use "power" and "energy" interchangeably.

packet transmission technique to further improve the delay performance. Detailed simulation studies are carried out to justify the effectiveness and efficiency of the proposed framework.

As a final note, the proposed DELAR can serve as a general framework in which various energy saving techniques such as power saving modes, transmission power control, power MAC and power-aware routing can be integrated to jointly achieve better energy conservation. In addition, it also offers a platform to study other challenging issues, e.g., quality of service (QoS) provisioning and security support. For instance, P-nodes can act as distributed admission controllers to coordinate the access to limited network resources such as available bandwidth. In security-sensitive applications, P-nodes can help B-nodes perform resource-hungry public-key operations.

The rest of the paper is organized as follows. We start with the review of related work in Section II. We then introduce the system model and the overall framework of DELAR in Section III. In Section IV, we elaborate on the network layer components of DELAR and a hybrid transmission scheduling scheme, and present a novel asymmetric MAC protocol called *A-MAC*, followed by a multi-packet transmission scheme to improve the delay performance. In Section V, we evaluate the performance of DELAR through simulations. Finally, we summarize this paper in Section VI.

II. Related Work

Recent years have witnessed a growing body of research concerned with energy conservation in mobile ad hoc networks, among which many efforts have been made at the physical layer to improve the hardware design of mobile devices [1]. Though important, it is still pertinent to explore other venues to further ameliorate the energy efficiency of mobile devices. For the lack of space, here we only review works on energy saving at the MAC or network layer, which are closely related to this paper. Based on the mechanisms used, they can be roughly classified into three categories: Power-Saving Modes (PSM), Transmission Power Control (TPC) [4], and Power-Aware Routing (PAR).

Coupled with scheduling, PSM is usually implemented at the MAC layer. The basic idea of PSM is to put the network interface into the sleep mode when no communication is needed. One fundamental issue in PSM is when to enter the sleep mode and when to wake up. Some work along this line includes PAMAS [5] and S-MAC [6]. In addition, Tseng *et al.* proposed three asynchronous protocols, namely, Dominating-Awake-Interval, Quorum-based, and Periodical-Fully-Awake-Interval protocols [7]. Another asynchronous scheme is reported in [8]. These proposals strive to efficiently and intelligently control nodes' sleep and wake schedules and at the same time deal with factors such as clock synchronization, neighbor discovery, and network partitions which are inherent in multi-hop ad hoc networks [7].

TPC adapts the transmission power to the channel and interference characteristics experienced by the link [9]–[17]. TPC sometimes is called topology control when it attempts to control the transmission power or even turn off the radio so that a desirable topology or connectivity can be maintained to save energy. Many proposals in this category are concerned with maintaining a dominant set of nodes or forming some virtual backbone with certain clustering mechanisms. In addition, PCM [18] was proposed to use different transmission power levels for RTS/CTS and DATA/ACK frames on a per-packet basis.

In contrast to TPC protocols aiming at making each link as energy-efficient as possible, a PAR protocol determines which of these links to be used for end-to-end paths so that additional energy savings can be obtained by routing packets over energy-efficient paths. The path optimality heavily relies on the routing metric employed in the routing protocol. Various factors could be considered in the design of the routing metric. For example, In [19], five metrics are discussed, for one of the optimization objectives: per-packet energy consumption, per-node lifetime and overall network lifetime. Besides, a single or a combination of perpetual or transient nodal or environmental characteristics can be used to construct a desired routing metric. Among those characteristics, node residual energy, transmission power level [20], link quality [21] (e.g., the cost for potential retransmissions). As an addition, PARO [22] is another notable approach designed for scenarios where nodes can dynamically adjust their transmission power. In PARO, a candidate intermediate node monitors an ongoing direct communication between two nodes and inserts itself in the forwarding path if its action can lead to some energy savings.

Most previous proposals have not considered device heterogeneity inherent in MANETs to achieve better energy conservation. How to take full advantage of P-nodes to prolong the network lifetime as much as possible has been previously addressed in [2], [23]–[27]. However, most of them focus on the network layer and many challenging issues related to the MAC layer are either left untouched or overlooked. For example, none of them considers how to support the MAC-layer acknowledgements over unidirectional links caused by different transmission power at P-nodes and B-nodes (cf. Section IV-D). In [25], a tunneling approach is used to convey the long-delayed MAC layer control packets via network layer over those unidirectional links (e.g., from B-nodes to P-nodes). The absence of timely MAC-layer acknowledgement still pose challenges to MAC operations as well as delay-sensitive applications. In [28], the authors proposed a MAC protocol for networks with asymmetric link, which is similar to our asymmetric MAC used in [29]. In our prior works [24], [29], [30], we did take the heterogeneity into consideration. In particular, the original idea of DELAR has already been presented in [29] and this paper provides the complete work.

In summary, although all the aforementioned schemes can achieve certain level of energy conservation, how to design a comprehensive, practical framework that not only integrates PSM, TPC, PAR, and transmission scheduling, but also makes the best use of inherent device heterogeneity remains an open, but challenging problem. Our DELAR framework is proposed to address this important problem.

III. Overview of DELAR

In this section, we present a high-level overview of the *DELAR*: Device-Energy-Load Aware Relaying framework for heterogeneous ad hoc networks.

A. Problem Statement

In this paper, we focus on heterogeneous ad hoc networks comprising of mobile nodes with different energy supplies, though heterogeneity in other forms may be possible [3]. We assume that besides a majority of battery-powered B-nodes, there also exist some powerful P-nodes having relatively unlimited energy supplies (cf. Section I). Our objective is to develop energy conservation protocols by utilizing such heterogeneity in energy resources. Intuitively speaking, since P-nodes are of "relatively" infinite energy reservoir as opposed to the B-nodes with usually irreplaceable batteries, data forwarding should attempt to utilize these P-nodes as much as possible in order to conserve energy of B-nodes. Therefore, on the one hand, a packet should be forwarded to a P-node whenever an energy saving can be expected. On the other hand, communications in the networks should avoid using B-nodes if possible. For these purposes, it is desirable to allow P-nodes to have higher transmission power so as to cover a larger transmission area, which can statistically reduce the number of B-nodes involved in packet forwarding. However, this straightforward proposal may pose significant design challenges–How can a B-node be aware of the existence of P-nodes in its vicinity? If there exist multiple paths through P-nodes to the destination, which path should be chosen? Should the transmission range of a P-node be as large as possible, or be kept within certain "optimal" ranges? How can the protocol support reliable communications over the unidirectional links caused by asymmetric transmission ranges between P-nodes and B-nodes along with error-prone and time-varying wireless channels? In addition, higher transmission power often implies more reachable neighbors, decreased spatial reuse, and increased local contention for the shared wireless medium, then how can the protocol schedule the transmissions so that a good balance can be struck between energy saving and other network performance factors such as packet delivery ratio and end-to-end delay? These are all non-trivial questions and need to be answered before we can indeed make the best use of the aforementioned heterogeneity in ad hoc networks.

After a careful examination on these interwoven issues, we believe that they are closely related to routing, transmission scheduling/MAC design, and power control. For example, with the adjustment of the transmission power of P-nodes and B-nodes, the topology and neighborhood, and thus the routing information, would change accordingly, and so would the schedule of transmissions. Moreover, there exists a strong interaction between routing and MAC layers. Apparently, our original design objective can be boiled down to designing a joint routing, scheduling, and power control scheme, which should be addressed across the whole protocol stack, especially at the routing and MAC layers [31] [32]. To achieve this, a cross-layer designed framework is demanded. In this framework, power control should be implemented to optimize the transmission power of each node (both P-nodes and B-nodes) to achieve optimal energy utilization and maintain a reasonably stable network topology; routing should be designed to inform all the nodes of the existence of P-nodes and find the optimal energy-efficient routes; and transmission scheduling should be able to adjust the transmissions so that the energy expended on channel contentions and collisions can be minimized. In addition, an appropriate scheduling scheme should be capable of striking a good balance between energy efficiency and other network performance metrics such as end-to-end delay and packet delivery ratio.

B. Our Solution: DELAR

Consider a mobile ad hoc network consisting of N_p P-nodes and N_b B-nodes, where N_p and N_b are system design parameters. We assume a single wireless broadcast channel shared by all the nodes, though our DELAR can be easily extended to multi-channel cases. We also adopt a simple power control scheme as follows. Each B-node transmits omni-directionally and can maintain a circular transmission range BTR (basic transmission range) before using up its battery, which can be properly set to achieve a good tradeoff between energy efficiency and network connectivity [33]. In addition, we postulate that P-nodes are able to adjust their transmission power so as to cover larger areas than B-nodes if needed. Moreover, all the P-nodes are assumed to have identical maximum transmission range of $PTR_{max} = M \times BTR$, where M is a positive integer greater than 1. As revealed in [33], using common transmission power between the same type of nodes can ensure bidirectional links and thus the correct operations of existing routing and MAC protocols. With this simple yet efficient power control scheme, a unidirectional link only exists between a P-node and a B-node when they use different transmission power, instead of between any two B-nodes or P-nodes[2]. According to [15], such simple power control is believed to be more practical than other expensive transmission power control schemes, either making unrealistic assumptions or having extra hardware requirements.

As mentioned before, DELAR arises from the following intuition: the P-nodes should be utilized as much as possible. In other words, we should attempt to minimize the use of B-nodes if possible. Thus it is advantageous to enable a P-node to directly communicate with other P-nodes nearby or at distance by using higher transmission power so that the number of B-nodes involved in the data forwarding can be reduced. However, higher transmission power or larger transmission coverage usually implies more neighbors, increased local contention and higher interference for the shared wireless channel. Therefore, instead of granting these P-nodes unlimited privileges of reaching any other node at any time at will, it makes more sense to constrain P-nodes' transmission power and control their transmissions in certain pre-planned manner in order to avoid the collisions with or minimize the interference to other ongoing transmissions.

In order to better schedule the transmissions of P-nodes and B-nodes, we adopt a time-division multiplexing method. We divide time into time slots of equal length, called *Superframes*. In each of the superframes, some time periods are exclusively designated to P-nodes, while the rest are shared by all P-nodes and B-nodes in the network. More specifically, during one cycle of the *Superframe* (see Fig. 1), there is a *P-to-P* period with length t_{pp}, in which only P-nodes are allowed

[2]In this paper, we only consider asymmetric transmission power as the primary cause for unidirectional links and omit others such as various collision/noise/interference levels at different nodes.

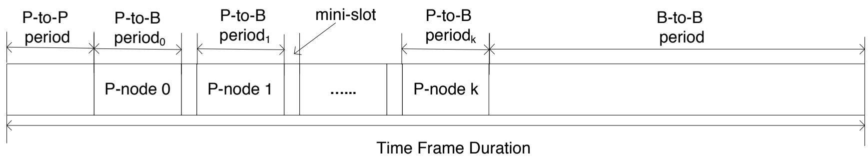


Fig. 1. The structure of a superframe.

to communicate with each other by using transmission range $TR_{pp} = m \times BTR$ $(1 < m \leq M)$, while all B-nodes just keep silent, as if the network were merely formed by these "mobile core" P-nodes. Additionally, in one *Superframe*, each P-node has its own exclusive period called *P-to-B* period of equal length t_{pb}, in which it can boost its transmission power to cover a range of $TR_{pb} = n \times BTR$ $(1 < n \leq M)^3$. The rest of one *Superframe* is called *B-to-B* period with length t_{bb} in which all the nodes in the network can contend for the channel and initiate transmissions towards nodes in their $TR_{bb} = BTR$. Obviously, all the P-nodes should act as common B-nodes in the *B-to-B* period by adjusting their transmission range back to TR_{bb}. Notice that during one *P-to-B* period, since the P-node owning this period and the B-nodes it intends to communicate with have different transmission power, unidirectional links between them may be formed. Therefore, in contrast to the *P-to-P* and *B-to-B* periods where some common contention-based MAC protocols such as the IEEE 802.11 can be used, the *P-to-B* period(s) demands an enhanced MAC protocol to support reliable communications over unidirectional links. Our Asymmetric MAC protocol *A-MAC* is developed exactly for this purpose. Such rendezvous of reservation-based and contention-based MAC schemes enables us to schedule the packet transmission more efficiently, which will be seen shortly.

In DELAR, the heterogeneity of mobile nodes is also incorporated into the construction of routing tables. Routes are discovered based on routing metrics which take the residual energy and the load status into consideration. Once generating a data packet, a node looks up its routing table and sends the data packet to the next hop as it does in common ad hoc routing protocols. If residing in the forwarding path and having received a forwarding request, a node will forward the data in an appropriate time period to the next hop according to its own routing table. More specifically, for a B-node, when the next hop is in its TR_{bb} range, it can only forward the data packet during the *B-to-B* period. While for a P-node, if the next hop is another P-node located in this P-node's TR_{pp}, the P-node can forward the packet to the next hop in the *P-to-P* period; if the next hop is a B-node located inside its TR_{pb}, the P-node can forward the packet to the next hop in its exclusive *P-to-B* period. In summary, with such time-division scheduling and a device-energy-load aware routing metric, we intend to utilize P-nodes as much as possible in an efficient and cautious manner while controlling the collisions and interference to an acceptable level in order to achieve the expected energy conservation without degrading the network performance.

[3]To provide reliable communications during *P-to-B* periods, usually n is less than m.

Several research challenges remain in supporting the seemingly simple operations of DELAR as described above. Given a B-node (P-node) X located in a P-node P's transmission range TR_{pb} (TR_{pp}), for instance, what criterion should P adopt to determine if X is a neighbor (in one hop range) or not, i.e., forwarding a packet to this node X in a one-hop manner or a multiple-hop manner? What kind of routing metric should be adopted to reflect the heterogeneity in device types, nodal residual energy, and local load status when setting up routing paths? How can we divide time into *Superframes*, and how can one P-node register a *P-to-B* period without conflicting with others' *P-to-B* periods? How does node X send MAC layer acknowledgements back to P in the presence of unidirectional links due to the asymmetry in transmission power? The remainder of this paper will address these questions one by one in more detail.

IV. Detailed Design of DELAR

In this section, we will first discuss the neighbor-selection criterion of P-nodes followed by the routing component of DELAR. We then introduce the detailed hybrid transmission scheduling of DELAR. Next, we present the Asymmetric Media Access Control Protocol (A-MAC) and the multi-packet transmission scheme. Last we give some further discussions.

A. P-nodes' Neighbor Discovery

In the literature, two nodes are usually considered as neighboring nodes of each other when they are one hop away and they can directly communicate with each other. However, in heterogeneous networks, we have to change the criterion to cope with the existence of P-nodes whose transmission ranges are much larger than those of B-nodes. In this case, any node in a P-node's TR_{pb} could be a *neighbor candidate* of the P-node[4]. Nevertheless, in order to support the MAC layer acknowledgements, not all the candidates can be finally chosen as *neighbors* or be next hops in the routing table. Before presenting the rules that guide P-nodes to make selection decision, we first introduce the notions of *Forward Path* and *Backward Path*. For any node pair s and t, a *Forward Path* indicates the path derived from normal routing tables. For example, the *Forward Path*$(s,\ t)$ can be represented as $s \rightarrow N_1 \rightarrow ... \rightarrow N_k \rightarrow t$, where $\{N_i\}$ $(1 \leq i \leq k)$ denote the k intermediate nodes between s and t. For a given P-node P and any B-node X located in P's transmission range TR_{pb}, the *Backward Path*$(P,\ X)$ is defined as the *minimum-hop Forward Path*$(X,\ P)$ when all the nodes have a transmission range of BTR. It is worth noting that the *minimum-hop Forward Path*(X, P) is not necessarily the same as the

[4]However, due to the asymmetric transmission power, the discussed P-node may not be a neighbor of an individual node of those neighbor candidates.

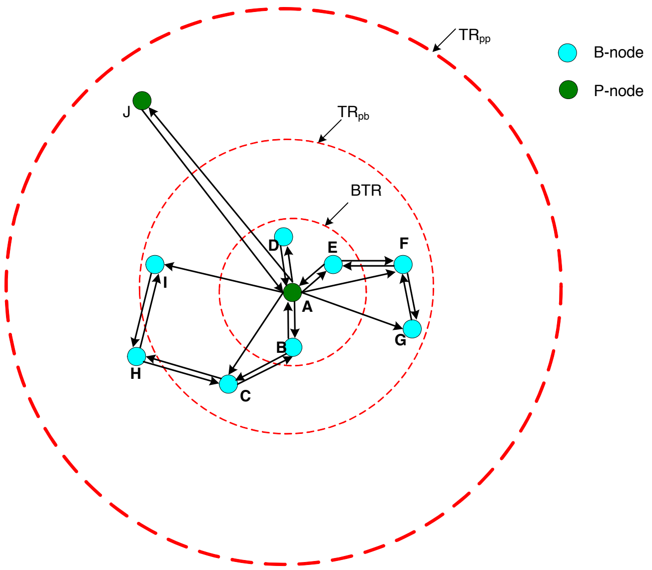


Fig. 2. An example of the neighbor selection process (m=4, n=2, T=3.).

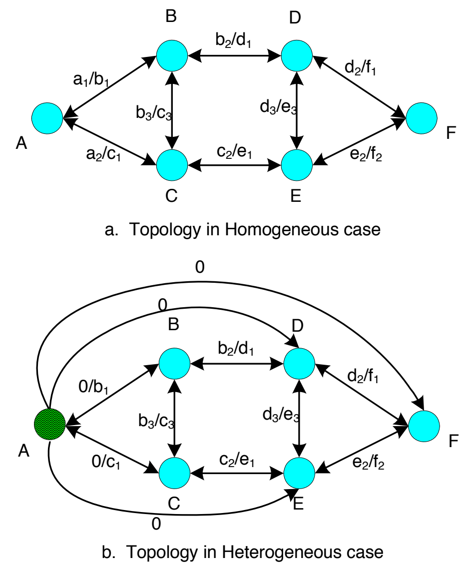


Fig. 3. The topology in homogeneous and heterogeneous cases.

Forward Path(X, P). Although *Forward Paths* are defined for any node pairs in the network, *Backward Paths* are only valid between a P-node and the B-nodes located within the P-node's TR_{pb} range. Furthermore, for any *neighbor candidate* X of a given P-node *P*, this B-node X can be considered as P's *neighbor* only when the *Backward Path(*P*,* X*)* satisfies the following criteria: All the intermediate nodes along the *Backward Path(*P*,* X*)* should be in P's TR_{pb} range. In other words, a *neighbor candidate* X can be considered as a P-node P's *neighbor* if and only if all the intermediate nodes along the *Backward Path(*P*,* X*)* are P's *neighbors* as well.

With the above definitions, the remaining issue is how to set up these *Backward Paths*. A simple way is to let a P-node broadcast a query message with a certain transmission power, i.e., covering all the B-nodes in its $TR_{pb} = n \times BTR$ range. Once seeing such a query, each node broadcasts a special reply with the TTL value set to T [5]. Each node appends its own ID in the reply when relaying such a special reply. The querying P-node will wait some time until collecting enough replies. The initiator of a reply would be considered as a neighbor if and only if the querying P-node also receives replies initiated from all the relaying nodes of that reply. We will see later, in order to facilitate the operation of A-MAC, the path length of a *Backward Path* should be limited. We need to point out that, even when a P-node, say P_1, receives a query message initiated by another P-node, say P_2, P_1 should reply like common B-nodes with a transmission range of *BTR*. Since our scheme is targeted for networks with low or moderate mobility, P-nodes can execute this process infrequently or in their respective *P-to-B* periods when topology changes are detected by the MAC protocol. Therefore, the resulting overhead is not significant.

P-nodes also need to discover the neighbor relationship among themselves. To achieve this, during the *P-to-P* period a P-node may send a query with appropriate transmission power that is set to cover a range of $TR_{pp} = m \times BTR$. P-nodes receiving this query may directly send replies back to the requesting P-node.

Fig. 2 gives an example of the neighbor selection process. Suppose A is a P-node with $TR_{pb} = 2 \times BTR$ and $TR_{pp} = 4 \times BTR$, and the Backward Paths for *neighbor candidates* C, F, G and I are $C \rightarrow B \rightarrow A$, $F \rightarrow E \rightarrow A$, $G \rightarrow F \rightarrow E \rightarrow A$, and $I \rightarrow H \rightarrow C \rightarrow B \rightarrow A$, respectively. Since node H does not initiate an reply to A, only C, D, and G are considered as A's *neighbors*. Of course, B, D, and E are A's *neighbors* as well. In this example, another P-node J is also a neighbor of node A because J is in A's TR_{pp} and they can directly communicate with each other.

B. Routing Component of DELAR

In traditional homogeneous ad hoc networks, a node can only communicate with other nodes in its BTR range, while in heterogeneous ad hoc networks, a P-node is able to reach any other node within a larger transmission range, e.g., TR_{pb} and TR_{pp}. Therefore, the resulting topology and routing strategy may be quite different from that in homogeneous networks. As an example, a network topology without P-nodes is depicted in Fig. 3.a, where all the links are bi-directional and labeled with equal or unequal costs on both directions. For instance, a_1/b_1 indicates that the link cost from A to B is a_1 while b_1 from B to A. In contrast, if one node, say, A, is identified as a P-node who can reach much further in the network, more unidirectional links may be added as shown in Fig. 3.b. We label unidirectional links from P-node A to its neighbors with cost 0 to represent node A's "unlimited" power supply.

To cope with such heterogeneity, each P-node needs to maintain an internal *neighbor table* recording its chosen neighbors within TR_{pb} and the corresponding *Backward Paths* of those neighbors. In addition, each node in the network, either a P-node or a B-node, needs to maintain a *forwarding routing table* similar to that in a normal table-driven routing protocol such as DSDV [34]. For each node *i*, we define

$$\beta(i) = residual\_energy(i) - \mu \times queue\_len(i),$$

[5] T can assume an integer value slightly larger than n to allow more replies.

where $residual\_energy(i)$ indicates current remaining energy level at node i, $queue\_len(i)$ represents the current load status at node i, and μ is a parameter representing the energy consumption per unit data transmission[6]. Then, the device-energy-load aware routing cost metric we adopt is given in Eq. 1, though other cost metrics are applicable in DELAR as well.

$$cost(i) = \begin{cases} 1/\beta(i), & \beta(i) > \gamma \\ a, & \beta(i) \leq \gamma \end{cases} \tag{1}$$

In the above cost metric, $cost(i)$ is the cost of using node i as a relay, it could be used as the cost of all directional links (arcs) starting from node i and directed to any of its neighbors; γ is a parameter used to adjust the weight of the awareness of load and energy in the overall cost metric; constant a assumes a relatively large value to avoid using the nodes short of energy. In addition, to represent a P-node's unique power capability or device type, instead of using Eq. 1, a P-node assumes a zero link cost[7] for all the links toward its B-node neighbors within TR_{pb} or P-node neighbors within TR_{pp}. Ideally, in order to find energy-efficient paths, each node should be informed about the routing costs of other nodes as accurate and prompt as possible which may lead to excessive overhead. In practice, however, the employed routing protocol should strike a good balance between energy efficiency and overhead. Proactive routing protocols are known for their capability of propagating network conditions through the whole network in due course so that appropriate QoS decisions, e.g., admission control and route selection, can be made intelligently. Thus, we adopt a proactive routing protocol, e.g., DSDV, as the underlying routing protocol to periodically exchange the routing information. We note that other types of routing protocols can also be used in this framework. After gathering enough routing information, a node can employ a shortest path algorithm to find the shortest paths and the related costs to all the other nodes in the network. Here the path cost is actually the sum of the cost defined in Eq. 1 of all the B-nodes along a forwarding path.

Similar to those energy-aware cost metrics proposed in the literature, by choosing the proper values of α, μ and γ, the cost function defined in Eq. 1 can help prolong the network lifetime by distributing the traffic more evenly throughout the network, avoiding the overuse of a small set of nodes, and consuming nodal energy resources in a more balanced manner [19], [35]. Moreover, DELAR spontaneously incorporates P-nodes' unique power capabilities or device types, residual energy information, and local load status into the routing protocol without using redirect tables in DEAR [24] any more.

C. Hybrid Transmission Scheduling

In order to mitigate the interference a P-node's transmission may cause to the ongoing transmissions, it is reasonable to only allow a P-node to boost its transmission power during some exclusively reserved periods. For this purpose, as we mentioned before, time is divided into time periods of equal length, called *Superframes*. Fig. 1 gives an instance of such a *Superframe* structure consisting of multiple reserved *P-to-B* periods, one for each P-node. The one-minislot-length paddings between consecutive *P-to-B* periods are used to further mitigate the possible interference. The period allocation of the *Superframe* can be designed as follows. During the network startup phase, P-nodes use high transmission power to communicate and negotiate with each other, determining the lengths of the *P-to-P* period, the *P-to-B* period and the *B-to-B* period, also associating each P-node with a *P-to-B* period. After finishing the negotiation, the P-nodes broadcast such allocation information to all the B-nodes in their own TR_{pb}. In this way, ultimately all the nodes are informed about the *Superframe* allocation, and can synchronize to such allocation[8]. The lengths of three periods mentioned above are determined as follows:

- The length of P-to-P period is determined by the number of neighboring P-nodes. If there are k neighboring P-nodes, then the length of P-to-P period $t_{pp} = l \cdot k$, where l is a system parameter.
- The total length of P-to-B periods is determined by the number of neighboring P-nodes and B-node distribution around each P-node. Every P-node will have its own P-to-B period (cf. Fig. 1). The length of each P-to-B period depends on the maximum number of hops of backward paths between that P-node and its neighboring B-nodes. Suppose the maximum number of hops of backward paths between P-node i and its neighboring B-nodes is m_i, then the length of its P-to-B period $t_{pb}(i) = d \cdot m_i$ $(1 \leq i \leq k)$, where d is a system parameter.
- The length of B-to-B period is determined by the number of neighboring B-nodes. If there are b neighboring B-nodes, then the length of B-to-B period $t_{bb} = q \cdot b$, where q is a system parameter.

In our design, a *P-to-B* period or *P-to-P* period can be shared among P-nodes far away from each other, if such sharing can ensure the transmissions conflict-free.

Since a P-node can communicate with other P-nodes in the *P-to-P* periods and communicate with the B-nodes within its BTR range in the *B-to-B* periods, it is natural that in its own *P-to-B* period, a P-node should give priority to packets intended to its B-node neighbors outside its *BTR* range but inside its TR_{pb} range. Thus, packet scheduling is needed at a P-node to determine the appropriate transmission schedule for the packets to be relayed or initiated by itself. In this paper we assume that nodes have perfect time synchronization and leave the synchronization problem in the networks as our future work.

As a remark, we notice that the time-division scheduling, essentially a reservation-based access control mechanism, and the MAC protocols employed in the three types of periods in each *Superframe*, either contention-based or reservation-based, form a hybrid transmission scheduling for DELAR. Moreover, each type of periods may use different MAC protocols. For example, the conventional contention-based MAC protocols,

[6]In our simulation, for example, μ is equal to the average energy consumption per packet transmission. Also note that μ is a function of time, i.e., it will be updated after a period of time or when the network topology is changed.

[7]In practice, a very small value can be used to avoid possible routing loops.

[8]For simplicity, we assume a fixed allocation scheme is used in this paper, however, more adaptive allocation is possible when P-nodes periodically exchange local load information and negotiate a new allocation scheme during the *P-to-P* periods.

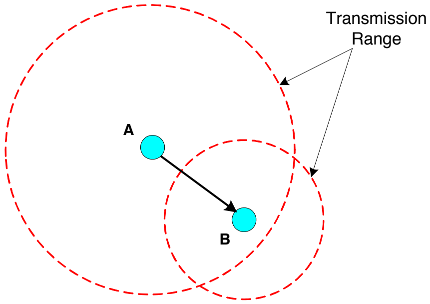


Fig. 4. A unidirectional link from A to B.

such as the IEEE 802.11 MAC protocol, can be used during *P-to-P* periods and *B-to-B* periods. Since unidirectional links are basically unavoidable in *P-to-B* periods, special measures are needed to deal with them. In what follows, we will delineate the A-MAC protocol developed for *P-to-B* periods.

D. Asymmetric Media Access Control Protocol (A-MAC)

The presence of unidirectional links is pretty common in heterogeneous networks especially when different transmission power levels are used. As an example, node *A* in Fig. 4 has a greater transmission range than node *B*. Thus, *B* can hear *A*'s transmission, however, *A* cannot detect *B*'s transmission, leading to a unidirectional link between *A* and *B*. The dilemma is that the stop-and-wait ARQ (Automatic Repeat Request) scheme [36] in current contention-based MAC protocols only works well with bidirectional links. In face of unidirectional links, the receiver *B* (Fig.4) has no way to directly and successfully send the acknowledgement back to the transmitter *A*, which may cause *A* to continuously transmit the same frame before timeout no matter whether *B* has received it or not. Moreover, the unidirectional links may severely affect the functionalities of ad hoc networks at various layers [37]–[39]. For example, many routing protocols such as DSR and AODV rely on hop-wise acknowledgments for discovering route errors. Therefore, how to support the MAC layer acknowledgements over the unidirectional links is very important [40], [41] and has not yet been well addressed. Fortunately, we can make use of the aforementioned *Backward Paths* and the following *"mini-routing"* method to tackle this problem in an elegant manner.

In current contention-based MAC protocols such as the IEEE 802.11, a receiver can only transmit an acknowledgement frame to its one-hop-away transmitter. With the cross-layer design methodology, we introduce a new concept of *"mini-routing"* into the MAC layer, which requests intermediate nodes to relay the receiver's acknowledgement frames, i.e., CTS/ACK frames, along the established *Backward Path(transmitter, receiver)* in a multi-hop fashion to the transmitter (a P-node) at the MAC layer. Here the routing information is no longer exclusively used by the network layer but also shared by the MAC and network layers.

Based on the IEEE 802.11 [42], we introduce into *A-MAC* four special frames: P-RTS, P-CTS, P-DATA, and P-ACK, all of which can only be transmitted in *P-to-B* periods. When a *P-to-B* period of a P-node comes and the P-node happens to have some packets to transmit, it first boosts its transmission power to cover the range of $TR_{pb} = n \times BTR$. With the scheduling described in Section IV-C, all the other nodes should refrain from initiating a transmission and temporarily suspend transmitting usual frames, i.e., RTS/CTS/DATA/ACK. The P-node associated with this *P-to-B* period can send packets to any neighboring B-node in the range of TR_{pb} through P-RTS/P-CTS/P-DATA/P-ACK exchanges.

Next, we illustrate the *A-MAC* by using Fig. 5, where we assume $n = 2$ and P-node A intends to send a packet to *C*, one of its B-node neighbors. The location relationship among *A*, *B*, and *C* is also depicted in Fig. 5. First, A sends the P-RTS with $TR_{pb} = 2 \times BTR$ containing the *Backward Path(A,C)* ($C \rightarrow B \rightarrow A$). Then according to the length of the *Backward Path(A,C)* in this example which is 2, A sets its waiting timer for the P-CTS to be $2(SIFS+T_{P-CTS}+T_{prop})$ [9], where T_{P-CTS} and T_{prop} are the transmission time and propagation time for one P-CTS, respectively. Upon receiving the P-RTS, and waiting for a *SIFS* period, node *C* will send node *B* a P-CTS containing the addresses of *A* and *C*. For an intermediate node residing on the backward path, it needs to set its waiting timer according to its order in the *Backward Path*. For example, the ith node (the intended receiver like *C* is assumed to be the *0*th node on the path) on the path should set its timer to be $i \times (SIFS + T_{P-CTS} + T_{prop})$. In this example, *B*, upon overhearing the above P-RTS from *A*, starts a timer equal to $SIFS+T_{P-CTS}+T_{prop}$ because it is the 1st intermediate node on the backward path. Once receiving a P-CTS from *C* before timeout, *B* simply appends its address and relays the modified P-CTS to the next hop which is the P-node *A* in this example. Otherwise, *B* sends a P-CTS containing its own address to *A* after the timer expires, and the reason for doing this will be explained later. If *A* does not receive any P-CTS before timeout, it can retransmit the P-RTS until reaching an admissible number of retries. If the same situation happens, *A* temporarily saves this packet for future transmission and switches to another packet with a different next-hop. When *A* successfully receives a P-CTS from *B* containing both *B's* and *C's* addresses, the P-RTS/P-CTS exchange finishes. After a *SIFS*, *A* can send a P-DATA frame to node *C* and set the timer to $2(SIFS + T_{P-ACK} + T_{prop})$. Then the similar procedures apply. After receiving the P-ACK from *C* relayed by the intermediate node *B*, the P-node *A* can start transmitting a new packet after a *DIFS* in the same manner. When its *P-to-B* period expires, *A* lowers its transmission power and acts as a B-node in other P-nodes' *P-to-B* periods and in the *B-to-B* period.

Besides the purpose of resolving the well-known hidden/exposed terminal problems, the P-RTS/P-CTS exchange is also used to eliminate possible errors resulting from stale routes or nodes' mobility. For example, in the above example, if C moves out of P-node A's $2 \times BTR$ range while *B* is still in *A*'s BTR range, node C will not hear the P-RTS from node A and hence *A* could only receive from *B* a P-CTS including only *B*'s address. In this case, *A* will think that *C* is currently unreachable and may temporarily save the packets to *C* for future transmissions. Another situation may happen that C is

[9]SIFS stands for Short Inter-frame Space, and DIFS stands for DCF Inter-frame Space in IEEE 802.11 standard.

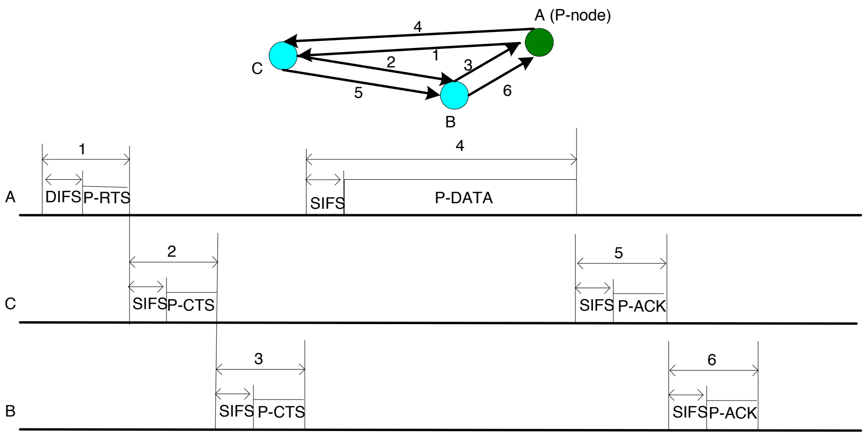


Fig. 5. The A-MAC operation procedure.

still in A's $2 \times BTR$ range while *B* moves out of *A's* BTR range, in which the P-node *A* will delete *C* from its neighbor table. Moreover, when *C* moves into *A*'s *BTR* range, *A* will hear the P-CTS from *C* directly. Hence *A* can optimize the future transmissions to *C* without the help of *B* any more.

As a remark, we notice that the MAC protocol in [28] bears some similarity to our A-MAC, but did not spell out how the acknowledgments will route back to the sender. Our A-MAC maintains the $Backward\ Path$ for such purpose, and hence should be more efficient.

E. The Multi-Packet Transmission Scheme

In the basic DELAR design, time-division transmission scheduling is used to coordinate the transmission activities of P-nodes and B-nodes. One undesirable consequence is the excessive delay one packet may experience because it may be buffered at intermediate nodes to wait for appropriate transmission periods. On the other hand, since DELAR is energy aware and it costs P-nodes almost "nothing" to transmit a packet, many data packets may swarm to P-nodes. This may make P-nodes the "bottlenecks" of the network and further increase the delay that packets accumulated at P-nodes would experience. In what follows, we seek a way to alleviate this phenomenon.

The basic idea is to enable multi-packet transmission with the help of A-MAC proposed in Section IV-D, which is illustrated using the topology in Fig. 6. During a *P-to-B* period, suppose P-node *A* intends to transmit some packets to both B-nodes *B* and *C*. In the original design, *A* can only transmit packets to either *B* or *C* each time. Moreover, node *C* has to rely on *B* to relay its acknowledgements to *A* because it is not within *A*'s BTR range. If multi-packet transmission is enabled, *A* would pack one packet for *C* and another packet for *B* together, and send them in a single packet from which nodes *B* and *C* can acquire their own part, respectively. In this way, we expect to see the improvement of the end-to-end delay performance of DELAR and potential throughput as well due to the elimination of the contention it may take when packet-by-packet transmissions are used.

To do this, *A* first makes sure that both B and C are within its effective transmission range TR_{pb} after the P-RTS/P-CTS exchange as before. Then A can bind one packet towards C with another packet towards B and send them in one P-DATA frame depicted in Fig. 6. When seeing such a frame, nodes *B* and *C* can extract their own packets and dump the rest. The same procedures as specified in A-MAC are executed with the exception that B also needs to indicate in the P-ACK that it has successfully received its packet.

One may expect that the total number of packets that can be packed and transmitted at one time is bounded by the length of the *Backward Path*. In fact, considering the possibility of adopting high-data rate modulation schemes with a higher power level so as not to degrade the received signal, more packets towards different receivers on the *Backward Path*, can be assembled together and transmitted at the same time. For lack of space, more results will be reported elsewhere.

F. Discussion

1) The existence of backward paths: In DELAR, the P-nodes are utilized in two ways to conserve energy: enabling P-nodes to directly communicate with other P-nodes within TR_{pp}, and enabling P-nodes to directly communicate with B-nodes within TR_{pb}. Unlike the communications between P-nodes that need no special treatment, the communications between P-nodes and B-nodes are supported by the nodes on the backward paths as described in Section IV-A, thus the backward path is very important for the proper functioning of DELAR. Then one may question the existence of such backward paths. We believe that the existence of backward paths is related to several factors such as basic transmission range (BTR), node density, and the location distribution of P/B-nodes. The backward path should not be too long in terms of hop count, otherwise A-MAC may not function efficiently because a long backward path is more subject to breakage due to node mobility. In practice, we should include this limitation into the aforementioned neighbor discovery. For example, the TTL value of a reply to the neighboring query can be set to n (cf. Section IV-A), such that only those B-nodes with backward paths of less than n hops can be considered as neighbors of a given P-node. When $TR_{pb} = 2 \times BTR$, the probability that there exists a two-hop backward path from a P-node to another node located in its TR_{pb} is pretty high under different network configurations. In fact, this probability

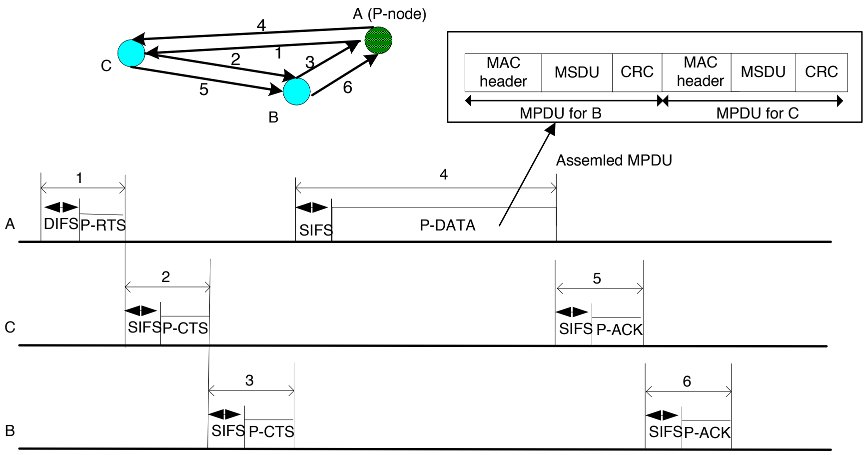


Fig. 6. The multi-packet transmission scheme.

will be higher when n increases. The existence of backward paths justifies the feasibility of our DELAR.

2) DELAR and ZRP: We can also borrow some ideas from zone routing protocol (ZRP) [43] to further improve the routing performance of DELAR as follows. A P-node maintains the routing information within a zone TR_{pb} by using the procedures described in Section IV-A or any other routing protocol as the intra-zone routing protocol (IARP). It also maintains the information about its neighboring P-nodes in its TR_{pp}. In contrast, a B-node only needs to maintain the routing information in its $TR_{bb} = BTR$, where all the nodes are its one-hop neighbors. Once a node needs an energy-efficient route to another node, it can discover the route on demand using the routing discovery procedure similar to that of AODV but with what is defined in Eq. 1 instead of hop count as the cost metric. In this sense, DELAR can be viewed as a special case of ZRP. The difference lies in the fact that all nodes within a P-node's TR_{pb} zone, are "one-hop" away from this P-node from the routing perspective rather than multi-hop away in the legacy ZRP. Besides, the border-casting technique used in ZRP [43], [44] is also available in DELAR in the sense that a P-node border-casts a route request (which we call simply a request) to all its peripheral nodes with corresponding backward paths embedded in the route request. In this way, each peripheral node would learn the backward path used to return a route reply if needed. Moreover, since a P-node can directly exchange routing information with other P-nodes within its TR_{pp}, it can also border-cast the route request to its P-node neighbors, which in turn can look up their own neighbor tables to decide if the desired destination is in their own TR_{pb} zones. Since TR_{pp} is usually larger than TR_{pb}, apparently, such "inter-zone routing protocol (IERP)" can further speed up the route discovery process.

3) The choice of m and n: Since a P-node can use higher transmission power to communicate with other P-nodes within its $m \times BTR$, larger m would lead to less use of B-nodes in the communications, but also less spatial reuse. Similarly, larger n may lead to more energy savings, but also imply possible longer backward paths and less spatial reuse, which may make A-MAC less efficient as explained before. For the similar reasons, usually n is less than m. In order to well balance the energy savings and other system performance factors, both m and n should be chosen appropriately. While DELAR requires all the P-nodes have the same value of m to avoid producing unidirectional links between P-nodes in *P-to-P* periods, with A-MAC in place, they can have different values of n.

4) Benefits of the time division scheduling: P-nodes can communicate with each other in one-hop or multi-hop manner during *P-to-P* periods to coordinate the use of the next *Superframe*. A P-node can also notify all the other nodes within its TR_{pb} to adjust their transmission schedules. With these intelligent communications available, the slot allocation in each *Superframe* can be adjusted adaptively according to traffic conditions instead of being fixed as in the given example. It is worth pointing out that this time-division scheduling method can facilitate the operations of PSM (cf. Section II) in that it can help nodes determine their sleep and wake schedules. For example, B-nodes can turn off the radios during the *P-to-P* periods to conserve energy.

As we discussed before, time synchronization is of importance for the correct operation of DELAR. In literature, there exist many proposals for time synchronization and many of them can be incorporated into DELAR. In fact, P-nodes can serve as coordinators to facilitate such synchronization by sending out some beacon information in some *P-to-P* periods and *P-to-B* periods periodically. Subsequently, B-nodes can synchronize their clocks to these P-nodes.

V. Performance Evaluation

In this section, we carry out the performance evaluation of our DELAR.

A. Simulation Setup

In order to evaluate the performance of DELAR, we implement our scheme including the routing layer and the *A-MAC* in the OPNET Modeler. We simulate a network with 50 nodes randomly deployed in a $1500 \times 300m^2$ area. The BTR is $200m$ and the transmission rate is $2Mbps$. In our simulations, all the nodes are capable of moving in the network according to the modified random Waypoint mobility model [45]. The pause time is set to be zero in our simulations, meaning nodes are always moving. Each node moves with a randomly chosen speed between [V_{min},V_{max}], where V_{min} is fixed to be 1 m/s and V_{max} assumes various values to reflect various network mobility levels.

There are 20 constant bit rate (CBR) data sessions between randomly selected source and destination pairs, and each source generates data packets of 512 bytes in length at a rate of λ packets per second. In our simulation, B-nodes have the same initial energy reservoir $3kJ$ and their transmission and reception powers are $1560mW$ and $930mW$, respectively. The networks with 2 to 6 P-nodes are studied and these P-nodes are randomly deployed. Besides, we choose $m = 4$, $n = 2$, $t_{pp} = 0.05s$, $t_{pb} = 0.05s$, and $t_{bb} = 0.15s$. By varying the number of P-nodes, the maximum moving speed, and the CBR source rate, we are able to study the performance of DELAR under various configurations. Six runs are carried out to get an average result for each simulation configuration and each run is executed for 900 seconds of simulation time.

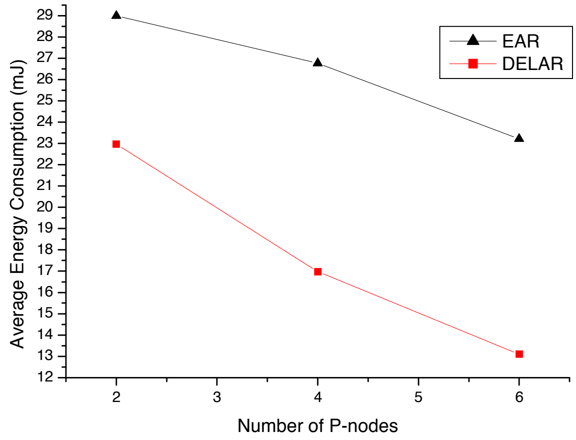
Previous work [19], [35] has shown that the energy efficiency of routing protocols can be much improved by adopting a routing metric such as what we defined in Eq. 1. Therefore, we shall only compare our DELAR with the one referred to as *EAR*, which is a variant of DSDV with the routing cost metric defined in Eq. 1. In EAR, all the P-nodes have the same transmission range as B-nodes, but have a zero cost to their neighbors because they are assumed to have almost unlimited energy reservoir. The following network performance metrics will be adopted in comparison: **average energy consumption** defined as the total energy consumption for all packet transmissions and receptions normalized by the number of delivered packets; **packet delivery ratio** defined as the ratio of delivered data packets to those generated by the sources; **average packet end-to-end delay**.

B. Impact of The Number of P-nodes

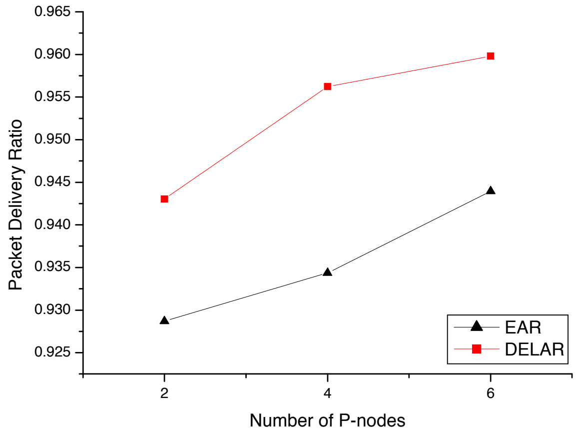
We first fix the data rate to 4 packets/s and vary the number of P-nodes to study its impact on the performance of DELAR. Here only the results for V_{max}=4 m/s are presented in Fig. 7, though DELAR has the similar performance with other values of V_{max}.

Fig. 7(a) compares the average energy consumption of DELAR and EAR under different numbers of P-nodes. Since EAR is also an energy-aware routing protocol, it is of no surprise to see that its energy-saving performance improves with the increase of the number of P-nodes whose routing costs to their neighboring nodes are assumed to be zero. With DELAR in place, however, the average energy consumption can be much further reduced, e.g., with a factor of almost 50% if 6 P-nodes are available. The displayed advantage comes from the fact that DELAR makes much better use of P-nodes than EAR through intelligent transmission scheduling and the allowance of P-nodes using different transmission power levels during various periods. The more P-nodes, the more energy savings we can expect from DELAR.

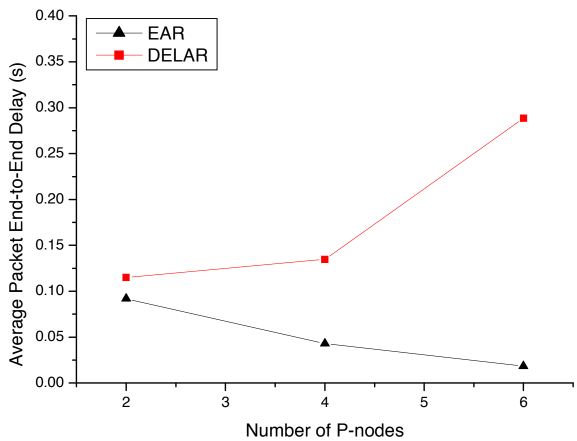
From Fig. 7(b), we can see that DELAR outperforms EAR in terms of packet delivery ratio (PDR). The reason is that transmission scheduling in DELAR will lead to fewer transmission collisions, and the larger transmission ranges of P-nodes during *P-to-P* periods and *P-to-B* periods can help reduce the number of hops a packet may travel. Again, the more P-nodes, the more PDR is improved.



(a) Average energy consumption



(b) Packet delivery ratio



(c) Average packet end-to-end delay

Fig. 7. Simulation results with different number of P-nodes.

In terms of end-to-end delay, however, EAR is better than DELAR as shown in Fig. 7(c). This is because, in contrast to the contention-based transmissions in EAR, DELAR divides the time into *Superframes* to schedule the transmission activities so that a packet usually needs to be buffered at a node waiting for the arrival of a proper transmission period. In addition, we can observe that the delay of EAR decreases with the increase of P-nodes because the existence of more P-nodes can achieve better load balance, that is, the number of energy-efficient paths may be increased. However, this is not always the case with DELAR. Besides the better load balance, the increase of P-nodes also means for DELAR that the delay from the transmission scheduling becomes larger. This contributes to the longer delay of DELAR.

C. Impact of Node Mobility

In this subsection, we study the impact of the node mobility on DELAR by varying V_{max} from 2 m/s to 16 m/s. For the reason of clarity, only the results for 4 P-nodes and a data rate of 4 $packets/s$ are presented.

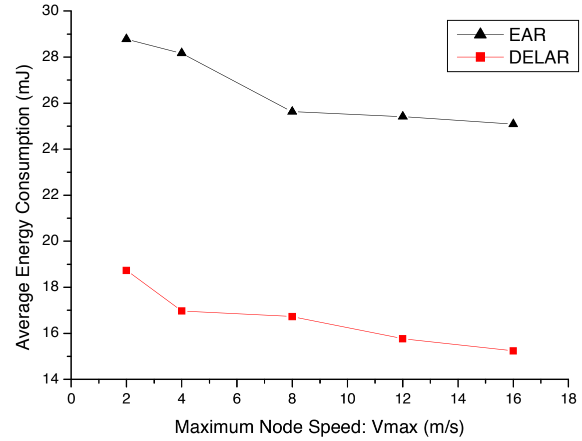
Fig. 8(a) compares the average energy consumption of DELAR and EAR under different mobility levels. As we can see, DELAR always has less energy consumption than EAR due to the reasons stated before. Generally the higher mobility leads to less energy consumption. After examining the average number of hops a packet may travel, we notice that higher mobility often results in shorter routes, which statistically leads to less energy consumption because fewer transmissions and receptions are involved. As shown in Fig. 8(b), the packet delivery ratio decreases with the increase of the mobility, which is in accordance with previous studies. For the similar reason we stated in the previous subsection, DELAR always has higher packet delivery ratio than EAR in all kinds of mobility. As the mobility increases, generally the delays of both EAR and DELAR get longer. Again, DELAR has longer delay than EAR due to DELAR's time-division medium access control mechanism. One interesting observation is that the delays of both DELAR and EAR fluctuate around $V_{max} = 2m/s$ and $V_{max} = 4m/s$. This can be attributed to the used routing cost metric which causes many packets swarming to the P-nodes. This phenomenon results in longer waiting time at P-nodes. However, nodal movement helps alleviate such phenomenon by dispensing the traffic load.

D. Impact of Traffic Load

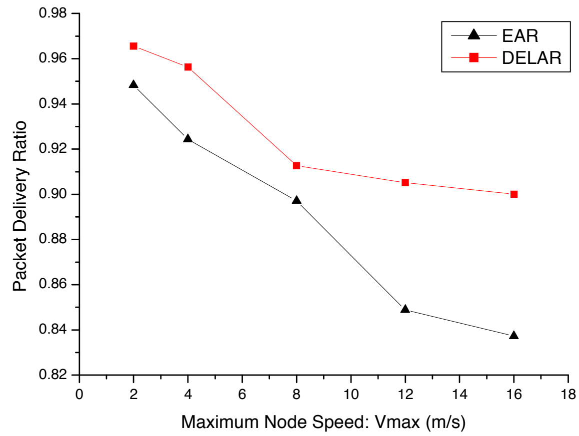
In this subsection, we study the impact of the traffic load on DELAR by varying the data generation rate from 3 $packets/s$ to 8 $packets/s$. Since the traffic load has no significant impact on the average energy consumption, we only depict the simulation results for the packet delivery ratio and the end-to-end delay in Fig. 8, where the number of P-nodes is four and V_{max} is equal to 4 m/s.

Fig. 9(a) and Fig. 9(b) demonstrate that the packet delivery ratio decreases and the delay increases with the increase of the traffic load for both schemes, which are quite intuitive. Again, our DELAR is better in terms of the packet delivery ratio, but worse with respect to the end-to-end delay than EAR for the reasons stated previously.

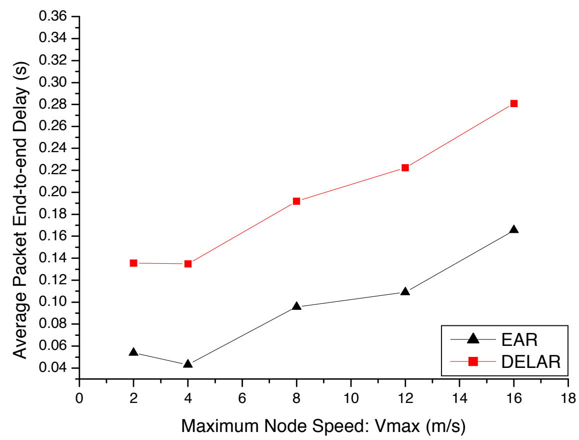
In summary, DELAR is more appropriate for delay-insensitive applications, such as file transfer and web access, which prefer higher energy-efficiency and packet delivery ratio.



(a) Average energy consumption

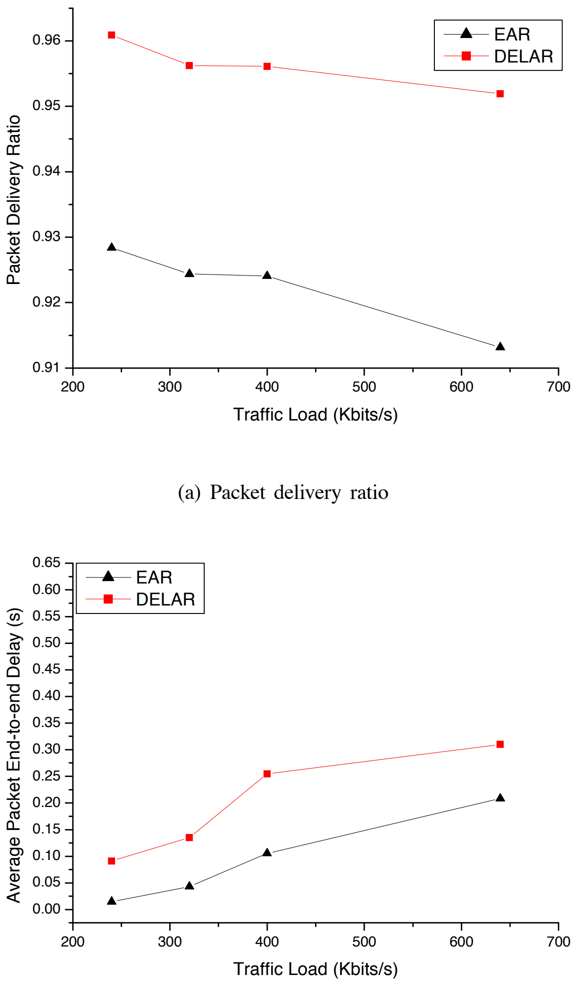


(b) Packet delivery ratio



(c) Average packet end-to-end delay

Fig. 8. Simulation results with different maximum node speed.



(a) Packet delivery ratio

(b) Average packet end-to-end delay

Fig. 9. Simulation results with various traffic load.

VI. Conclusions

In this paper, we propose a Device-Energy-Load Aware Relaying framework, namely DELAR, to achieve energy conservation in heterogeneous mobile ad hoc networks. DELAR utilizes the device heterogeneity inherent in ad hoc networks and features the cross-layer protocol design methodology. To take better advantage of powerful nodes (P-nodes) while mitigating their interference to the ongoing communications, a hybrid transmission scheduling mechanism is used to schedule and coordinate the transmission activities among P-nodes and B-nodes (normal nodes). In addition, in order to support reliable transmissions in the presence of unidirectional links between P-nodes and B-nodes, we introduce the "mini-routing" technique and the novel Asymmetric MAC (*A-MAC*) protocol, which demonstrates that *A-MAC* can effectively enable the MAC layer acknowledgements over unidirectional links. We have shown that DELAR can significantly reduce the energy consumption and thus prolong the network lifetime even with just a few P-nodes placed in the network. With this framework, various energy conservation techniques such as power saving modes, transmission power control and power-aware routing can be integrated to jointly achieve better energy conservation. More importantly, this framework provides a platform to address other challenging issues such as QoS provisioning and security support as well.

References

[1] C. Jones, K. S. P. Agrawal, and J.-C. Chen, "A survey of energy efficient network protocols for wireless networks," *ACM/Kluwer Journal on Wireless Networks (WINET)*, vol. 7, no. 4, August 2001.

[2] I. Chakeres and E. Belding-Royer, "Resource biased path selection in heterogeneous mobile networks," in *UCSB Technical Report 2003-18*, July 2003.

[3] M. Yarvis, A. Kushalnagar, H. Singh, Y. Liu, and S. Singh, "Exploiting heterogeneity in sensor networks," in *Proc. IEEE INFOCOM'05*, March 2005, pp. 878–890.

[4] M. Krunz, A. Muqattash, and S.-J. Lee, "Transmission power control in wireless ad hoc networks: challenges, solutions and open issues," *IEEE Network*, vol. 18, no. 5, pp. 8–14, 2004.

[5] S. Singh and C. S. Raghavendra, "PAMAS-power aware multi-access protocol with signaling for ad hoc networks," *Computer Communication Review*, vol. 28, no. 3, pp. 5–26, 1998.

[6] J. H. W. Ye and D. Estrin, "An energy-effcient mac protocol for wireless sensor networks," in *Proc. INFOCOM'02*, 2002.

[7] Y.-C. Tseng, C.-S. Hsu, and T.-Y. Hsieh, "Power-saving protocols for IEEE802.11-based multi-hop ad hoc networks," in *Proc. of INFOCOM'02*, 2002.

[8] J. Youn and C. Kang, "Asynchronous power saving schemes for ad hoc wireless networks," in *Information Technology Journal*, vol. 7, 2008, pp. 277–284.

[9] V. Roodoplu and T. Meng, "Minimum energy mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1333–1344, August 1999.

[10] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, "SPAN: an energy-efficient coordination algorithm for topology maintainence in ad hoc wireless networks," *ACM/Kluwer Journal on Wireless Networks (WINET)*, vol. 8, no. 5, Sept. 2002.

[11] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," in *Proc. MobiCom'01*, 2001.

[12] D. Blough, M. Leoncini, G. Resta, and P. Santi, "The k-NEIGH protocol for symmetric topology control in ad hoc networks," in *Proc. MobiHoc'03*, 2003.

[13] J. Gomez and A. T. Campbell, "A case for variable-range transmission power control in wireless ad hoc networks," in *Proc. IEEE INFOCOM'04*, Hongkong, China, March 2004.

[14] W.-Z. Song, Y. Wang, X.-Y. Li, and O. Frieder, "Localized algorithms for energy efficient topology in wireless ad hoc networks," in *Proc. MobiHoc'04*, 2004.

[15] A. Muqattash and M. Krunz, "POWMAC: A single-channel power-control protocol for throughput enhancement in wireless ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 5, pp. 1067–1084, 2005, special Issue on Advances in Military Wireless Communications.

[16] S. Zarifzadeh, A. Nayyeri, and N. Yazdani, "Efficient construction of network topology to conserve energy in wireless ad hoc networks," in *Computer Communications*, vol. 31, Jan. 2008, pp. 160–173.

[17] X. Huang and I. Rubin, "Bit-per-joule performance of power saving ad hoc networks with a mobile backbone," in *Wireless Networks*, vol. 16, 2010, pp. 311–329.

[18] E.-S. Jung and N. Vaidya, "A power control MAC protocol for ad hoc networks," in *Proc. MobiCom'02*, 2002.

[19] S. Singh, M. Woo, and C. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. MobiCom'98*, 1998.

[20] J.-H. Chang and L. Tassiulas, "Energy conservation routing in wireless ad-hoc networks," in *Proc. INFOCOM'00*, 2000.

[21] S. Banerjee and A. Misra, "Minimum energy paths for reliable communication in multi-hop wireless networks," in *Proc. MobiHoc'02*, 2002.

[22] J. Gomez, A. Campbell, M. Naghshineh, and C. Bisdikian, "PARO: Supporting dynamic power controlled routing in wireless ad hoc networks," *ACM/Kluwer Journal on Wireless Networks (WINET)*, vol. 9, no. 5, September 2003.

[23] J.-H. Ryu and D. Cho, "A new routing scheme concerning power-saving in mobile ad hoc networks," in *Proc. ICC'00*, 2000.

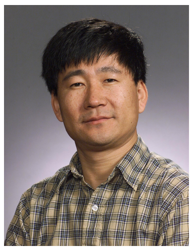
[24] A. Avudainayagam, W. Lou, and Y. Fang, "DEAR: A device and energy aware routing protocol for heterogeneous ad hoc networks," *Journal of Parallel and Distributed Computing*, vol. 63, no. 2, pp. 228–236, Feb. 2003.

[25] V. Shah, E. Gelal, and S. Krishnamurthy, "Handling asymmetry in power heterogeneous ad hoc networks," in *Computer Networks*, vol. 51, 2007, pp. 2594–2615.
[26] L. Song, A. Arora, and H. Zhang, "On exploiting asymmetric wireless links via one-way estimation," in *Proc. ACM MobiHoc'07*, 2007, pp. 11–21.
[27] V. Ramasubramanian and D. Mosse, "BRA: A bidirectional routing abstraction for asymmetric mobile ad hoc networks," in *IEEE/ACM Trans. Netw.*, vol. 6, 2008, pp. 116–129.
[28] G. Wang, D. Turgut, L. Bölöni, Y. Ji, and D. Marinescu, "A MAC layer protocol for wireless networks with asymmetric links," *Ad Hoc Networks*, vol. 6, no. 3, pp. 424–440, 2008.
[29] W. Liu, Y. Zhang, W. Lou, and Y. Fang, "DELAR: Device/Energy/Load aware relaying in heterogenous mobile ad hoc networks," in *Proc. MILCOM 2004*, Montery, CA, October 2004.
[30] X. Du, D. Wu, W. Liu, and Y. Fang, "Multiclass routing and medium access control for heterogeneous mobile ad hoc networks," *IEEE Trans. Veh, Technol.*, vol. 55, no. 1, pp. 270–277, 2006.
[31] U. Kozat, I. Koutsopoulos, and L. Tassiulas, "A framework for cross-layer design of energy-efficient communication with QoS provisioning in multi-hop wireless networks," in *Proc. IEEE INFOCOM'04*, Hongkong, China, March 2004.
[32] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: Joint routing, scheduling and power control," in *Proc. IEEE INFOCOM'04*, Hongkong, China, March 2004.
[33] V. Kawadia and P. R. Kumar, "Principles and protocols for power control in ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 5, pp. 76–88, January 2005, special Issue on Wireless Ad Hoc Networks.
[34] C. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *Proc. ACM SIGCOMM'94*, Sept. 1994, pp. 234–244.
[35] C. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE Commun. Mag.*, vol. 39, no. 6, pp. 138–147, June 2001.
[36] A. Leon-Garcia and I. Widjaja, *Computer Networks:Fundmental Concepts and Key Architectures*. McGraw-Hill, 1999.
[37] L. Bao and J. Garcia-Luna-Aceves, "Link-state routing in networks with unidirectional links," in *Proc. ICCCN'99*, 1999.
[38] M. Marina and S. Das, "Routing performance in the presence of unidirectional links in mulithop wireless networks," in *Proc. MobiHoc'02*, Lausanne, Switzerland, June 2002.
[39] R. Prakash, "A routing algorithm for wireless ad hoc networks with unidirectional links," *ACM/Kluwer Journal on Wireless Networks (WINET)*, vol. 7, pp. 617–625, 2001.
[40] S. Nesargi and R. Prakash, "A tunneling approach to routing with unidirectional links in mobile ad-hoc networks," in *Proc. ICCCN'00*, 2000.
[41] V. Ramasubramanian, R. Chandra, and D. Mosse, "Providing a bidirectional abstraction for unidirectional ad hoc networks," in *Proc. INFOCOM'03*, 2003.
[42] "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," IEEE 802.11WG, August 1999.
[43] M. Pearlman and Z. Haas, "Determining the optimal configuration for the zone routing protocol," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1395–1414, August 1999.
[44] Z. Haas, M. Pearlman, and P. Samar, "The zone routing protocol (ZRP) for ad hoc networks," IETF Internet Draft, draft-ietf-manet-zone-zrp-04.txt, July 2002.
[45] J. Yoon, M. Liu, and B. Noble, "Sound mobility models," in *Proc. ACM MobiCom'03*, San Diego, CA, September 2003.

Wei Liu (S'03-M'05) received his B.E. and M.E. in Electrical and Information Engineering from Huazhong University of Science and Technology, Wuhan, China, in 1998 and 2001. In August 2005, he received his Ph.D. in Electrical and Computer Engineering from University of Florida. Currently, he is a senior technical member with Scalable Network Technologies. His research interest includes cross-layer design, and communication protocols for mobile ad hoc networks, wireless sensor networks and cellular networks.

Chi Zhang (S'06) received the B.E. and M.E. degrees in Electrical and Information Engineering from Huazhong University of Science and Technology, Wuhan, China, in July 1999 and January 2002, respectively. He is working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at the University of Florida, Gainesville, Florida, USA. His research interests include wireless networks and network security.

Guoliang Yao (M'10) received the Ph.D. degree in electronic engineering from Southeast University, China, in 2010, graduated with M.E. in electronic engineering form Southeast University, China, in 2005 and received the B.S degree in physics from Hangzhou Normal University in 2001. He is currently a postdoctoral associate in the Department of Electrical and computer engineering, University of Florida. His current research interests are in wireless networks protocol and cross-layer design, communications system analysis, wireless network real testbed construction.



Yuguang Fang (S'92-M'97-SM'99-F'08) received a Ph.D. degree in Systems Engineering from Case Western Reserve University in January 1994 and a Ph.D. degree in Electrical Engineering from Boston University in May 1997. He was an assistant professor in the Department of Electrical and Computer Engineering at New Jersey Institute of Technology from July 1998 to May 2000. He then joined the Department of Electrical and Computer Engineering at University of Florida in May 2000 as an assistant professor, got an early promotion to an associate professor with tenure in August 2003 and to a full professor in August 2005. He holds a University of Florida Research Foundation (UFRF) Professorship from 2006 to 2009, a Changjiang Scholar Chair Professorship with Xidian University, Xi'an, China, from 2008 to 2011, and a Guest Chair Professorship with Tsinghua University, China, from 2009 to 2012. He has published over 300 papers in refereed professional journals and conferences. Dr. Fang received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002, and is the recipient of the Best Paper Award in IEEE International Conference on Network Protocols (ICNP) in 2006 and the recipient of the IEEE TCGN Best Paper Award in the IEEE High-Speed Networks Symposium, IEEE Globecom in 2002. He has also received a 2010-2011 UF Doctoral Dissertation Advisor/Mentoring Award and the 2009 UF College of Engineering Faculty Mentoring Award.

Dr. Fang is also active in professional activities. He is a Fellow of IEEE and a member of ACM. He is currently serving as the Editor-in-Chief for IEEE Wireless Communications (2009-present) and serves/served on several editorial boards of technical journals including IEEE Transactions on Mobile Computing (2003-2008, 2011-present), IEEE Transactions on Communications (2000-present), IEEE Transactions on Wireless Communications (2002-2009), IEEE Journal on Selected Areas in Communications (1999-2001), IEEE Wireless Communications Magazine (2003-2009) and ACM Wireless Networks (2001-present). He served on the Steering Committee for IEEE Transactions on Mobile Computing (2008-2010). He has been actively participating in professional conference organizations such as serving as the Technical Program Co-Chair for IEEE INOFOCOM'2014, the Steering Committee Co-Chair for QShine (2004-2008), the Technical Program Vice-Chair for IEEE INFOCOM'2005, the Technical Program Area Chair for IEEE INFOCOM (2009-2012), Technical Program Symposium Co-Chair for IEEE Globecom'2004, and a member of Technical Program Committee for IEEE INFOCOM (1998, 2000, 2003-2008).