PAPER

Priority-Based STDMA Scheduling Algorithm to Enhance Throughput and Fairness in Wireless Mesh Networks**

Nguyen H. TRAN†a), *Nonmember*, Choong Seon HONG†*b), *Member*, *and* Sungwon LEE†c), *Nonmember*

SUMMARY The aggregate throughput of wireless mesh networks (WMNs) can be significantly improved by equipping the mesh routers with multiple radios tuned to orthogonal channels. Not only the links using orthogonal channels can be activated at a time, but some links in the same channel also can be activated concurrently if the Signal-to-Interference-and-Noise Ratio (SINR) at their receivers is not lower than the threshold, which is the spatial-reuse characteristic. STDMA is considered as one of the medium access schemes that can exploit spatial reuse to improve network throughput. Past studies have shown that optimizing the performance of STDMA is NP-Hard. Therefore, we propose a STDMA-based scheduling algorithm that operates in a greedy fashion for WMNs. We show that the proposed algorithm enhances not only the throughput but also the fairness by capturing the essence of spatial-reuse approach of STDMA and giving medium access opportunities to each network element based on its priority. We furthermore validate our algorithm through theoretical analysis and extensive simulations and the results show that our algorithm can outperform state-of-the-art alternatives.

key words: *wireless mesh networks, STDMA, scheduling algorithm, fairness*

1. Introduction

Wireless mesh networks (WMNs) have emerged as a new, cost-effective network for the next generation wireless Internet [1]. In such networks, mesh routers which are stationary or less mobile nodes constitute the infrastructure backbone for clients whereas the mesh clients are the wireless terminals to which the WMN provides connectivity. Only a fraction of nodes have direct access to the Internet and serve as gateways. Mesh routers usually convey a large amount of traffic generated by mesh clients through gateways to the Internet. Energy conservation is not an issue of WMNs in comparison to ad hoc or sensor networks because mesh routers have an unlimited power capability. Instead, the main design concern is increasing the traffic carrying capacity of mesh routers. In order to satisfy the high traffic demands, mesh routers can be equipped with multiple radios operating on multiple channels. Based on the benefits of multi-radio multi-channel mesh routers, several recent works have focused on many typical problems of WMNs like channel assignment, routing, scheduling [2], [6], [9], [11], [17]. Due to the limited wireless channel capacity, the large number of clients and the emergence of real-time multimedia applications, designing an efficient scheduling scheme that can improve network throughput has been a critical issue of WMNs.

It has been acknowledged that CSMA/CA is widely applicable to mobile ad-hoc networks (MANETs) since it does not need to know any topology or clock information, which makes CSMA/CA very robust to any topology changes in MANETs. But with stationary or less mobile mesh routers, WMNs have no advantage from this property of CSMA/CA. Furthermore, due to its conservative mechanism with carrier sensing and collision avoidance characteristics, CSMA/CA is not considered as a good candidate for fulfilling the high-traffic demand of WMNs [4], [16].

Another access scheme that could be considered is Time Division Multiple Access (TDMA), where each transmission is reserved only to one time-slot. If a link has no packet to transmit in its time-slot, this slot is free, which is a waste of bandwidth. Therefore TDMA is very inefficient to meet the high-bandwidth requirement of WMNs [3]. One of the important factors to improve the network capacity is spatial reuse, the total number of concurrent transmissions that can be accommodated in the network. In the history, Spatial TDMA (STDMA) [15] is an access scheme that provides the concurrent transmissions as long as they do not interfere too much with each other. In STDMA, how to exploit the efficiency of the spatial reuse is relying on the scheduling algorithms designed to generate transmission schedule in every time-slot. But there is a number of difficulties in employing STDMA-based scheduling algorithm to wireless networks that has been identified so far. One of them is how to derive a STDMA-based algorithm in a distributed fashion that also takes into account the traffic distribution. Another problem is finding the optimal schedule length (i.e., the minimum-length schedule). Even in the most trivial scenario with unit traffic demand on each link, it was shown that a centralized STDMA scheduling problem is NP-Hard out of the complexity computation viewpoint [3], [14].

Moreover, as the majority of traffic is transferred to and from gateways, traffic flows will likely aggregate at the mesh routers close to the gateways. So without efficiently designed scheduling algorithm, mesh clients of border mesh routers could suffer the starvation of bandwidth. So, be-

Manuscript received January 18, 2010.
Manuscript revised August 19, 2010.
†The authors are with the Computer Engineering Department of Kyung Hee University, South Korea.
*Corresponding author.
**This work was partially supported by the IT R&D program of MKE/KEIT [2009-S-014-01, On the development of Sensing based Emotive Service Mobile Handheld Devices] and Kyung Hee University in 2010 (KHU-2010372).
a) E-mail: nguyenth@khu.ac.kr
b) E-mail: cshong@khu.ac.kr
c) E-mail: drsungwon@khu.ac.kr
DOI: 10.1587/transcom.E94.B.1355

sides targeting to improve the overall throughput of the system, a good scheduling scheme also has to account for the link-fairness provisioning problem to give mesh routers at the border of WMNs appropriate transmission probabilities [12], [13], [19].

Based on discussions above, in this paper we propose a partially centralized, greedy, and priority-based STDMA scheduling algorithm for WMNs. Our proposed scheme is partially centralized because the essence of the algorithm is divided into two phases. The first phase is centrally operated by gateways, and the second phase is interacted distributively between mesh routers, which would alleviate the burden on gateway. The reason for choosing greedy approach is to substitute the NP-Hard complexity of the STDMA optimal performance by a simpler one. Also, with the priority-based medium access characteristic, more opportunities are granted to links which have more priorities (e.g. links that are closer to gateways), which enables the link-fairness provision. We note that link-fairness is a unique property of WMNs due to its special topology, traffic distribution and high bandwidth demand [12], [13]. So link-fairness here could be confused with flow fairness. However, surprisingly our proposed algorithm also outperforms the GreedyPhysical [4], which is the most comparable with our scheme, and CSMA/CA in terms of flow fairness besides link-fairness and throughput in simulations even though flow-fairness was not one of our initial objectives.

The remainder of our paper is organized as follows. The next section is the related work. In Sect. 3, we formulate the system models. Next, we describe our proposed algorithm in Sect. 4 and present the algorithm analysis in Sect. 5. We evaluate the performance of our algorithm through simulations in Sect. 6. Finally, we conclude our work in Sect. 7.

2. Related Work

STDMA access scheme was first introduced in [15]. Afterwards, it is mostly investigated in MANETS, where the objective to achieve the optimal performance is formulated by computing a set of non-conflicting transmission schedule, which is an NP-Hard problem that is intractable computationally [3], [8], [10]. There is a limited number of work focusing on STDMA-based scheduling algorithm for WMNs [4], [19].

In [19], Salem et al. propose a scheduling mechanism that ensures per-client fairness by assigning transmission opportunities to the most loaded links in a STDMA fashion. Even though the ideas are attractive by computing the maximal number of the most loaded links in a non-conflicting schedule at every time-slot, their solution is impractical for a kind of next generation network with high traffic demand like WMNs due to the fact that their solution that uses clique enumeration is also an NP-hard problem.

Perhaps the work that is closest to our work is the GreedyPhysical algorithm of Brar et al. [4]. This STDMA-based scheduling algorithm is a centralized scheme that also utilizes greedy approach to replace the intractable complexity of STDMA optimal performance. However, GreedyPhysical treats all links in WMNs equally based on the load on each link when constructing a schedule. Without considering the traffic distribution characteristic of WMNs, where links closer to gateways have higher bandwidth demands, this algorithm was not aware about the link-fairness problem of WMNs. This could lead to the throughput degradation [12], [13].

Specifically, even though the work of Liu et al. [13] did not use any specific medium access scheme but instead assume only a given channel-access probability for their analysis, it is the first work that characterizes location-dependent throughput, delay and fairness properties of WMNs by using extremely classical queueing model techniques. Location-dependent fair throughput objective of this paper draws similarly the target we aim to. That is the location of each link should be accounted for when designing a scheduling algorithm for WMNs in order to provide the link-fairness and so on enhance the network throughput.

3. System Models

3.1 Network Model

We consider the backbone of WMN modeled by a *network graph* $G(N, L)$, where N is the set of $|N|$ nodes (mesh routers) and L is the set of $|L|$ bi-directional links. We assume that time is slotted, denoted by t, and that the packet length is normalized in order to be transmittable in a unit time-slot. We denote $Q_l(t)$ the queue length of link l at time-slot t. In the system, each node is equipped with one or more wireless interface cards or radios. We denote I the set of $|I|$ interface cards (radios) and C the set of $|C|$ orthogonal channels available in the network. Because of half-duplex wireless interface, each node can not transmit and receive data at a time on a particular channel. However, by using multiple radios and multiple channels, an interface of a node can transmit the data on one channel while another interface can receive data on a different channel. Suppose that the maximum hop count distance from the gateway is K, then our system is called K-tier WMN. An example of a 5-tier WMN is illustrated in Fig. 1 with 4 orthogonal channels and each node can be equipped up to 4 wireless interfaces.

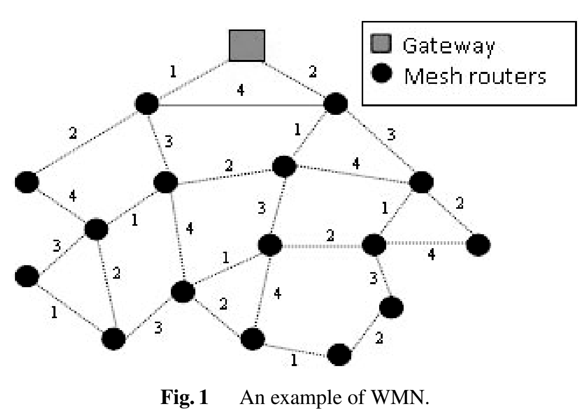


Fig. 1 An example of WMN.

For STDMA, there are two kinds of time-slot assignment: node-oriented and link-oriented assignment. Here, we choose link-oriented assignment for our proposed algorithm because empirically, link-oriented assignment can provide better spatial reuse than node-oriented assignment [3]; hence it satisfies the high traffic demand of WMNs. In link-oriented assignment, a link is allocated no less than one time-slot between a pair of mesh routers.

In order to schedule two links at the same time-slot, we must ensure that the schedule would be able to suffer the interference. In this model, a successful transmission from the transmitter $tx(l)$ to the receiver $rx(l)$ of link l depends on the SINR at $rx(l)$. Specifically, denoting $RSS^{tx(l)}_{rx(l)}$ the received signal strength at $rx(l)$, and $ISS^{tx(l_i)}_{rx(l)}$ the interfered signal strength received by $rx(l)$ from other transmitters $tx(l_i)$ of the set of links $\{l_i, i = 1, 2 \ldots\}$ which are also active simultaneously with l, packets along link l are correctly decoded if and only if:

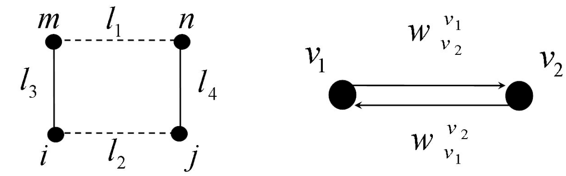
$$\frac{RSS^{tx(l)}_{rx(l)}}{W + \sum_i ISS^{tx(l_i)}_{rx(l)}} > \gamma_l \tag{1}$$

where W is the background noise and the constant value γ_l is the predefined threshold. Based on the interference model above, the set of communication links that interfere with each other can be represented by using interference graph.

3.2 Interference Graph

We denote an *interference graph* by $G^I(V, E)$, where vertices and edges are used to distinguish with nodes and links of network graph. The set of vertices V in interference graph corresponds to the communication links in the network graph and there is a directed edge of the set E connecting two vertices if there is interference between these corresponding two links in network graph. Each directed edge has its own weight. This weight represents what fraction of maximum permissible noise and interference level at the receiver of link l_2 (or l_1) (such that link l_2 (or l_1) is still operational) contributed by transmission on link l_1 (or l_2), denoted by $w^{v_1}_{v_2}$ (or $w^{v_2}_{v_1}$). The interference graph must be weighted because the more the links on the same channel are active concurrently, the more interference each link will be affected by each other until the packets received are incorrect. Consider a scenario in Fig. 2, the communications between node m and n, i and j are on the same channel (the dashed lines), for example channel 1. The communications between m and i, n and j are on channel 2 and 3 respectively. So there is interference between l_1 and l_2 and we can construct the interference graph based on the network graph as Fig. 2. From the above definition of weighted interference graph, we define the weight value $w^{v_1}_{v_2}$ which represents for the interference contributed by l_1 to l_2 (we can calculate $w^{v_2}_{v_1}$ as this way similarly):

$$w^{v_2}_{v_1} = \frac{ISS^{m}_{j}}{\frac{RSS^{i}_{j}}{\gamma_{l_2}} - W}, \tag{2}$$



(a) Network graph

(b) Weighted interference graph

Fig. 2 Network graph and weighted interference graph.

where $\frac{RSS^{i}_{j}}{\gamma_{l_2}} - W$ is the maximum permissible interference noise which still allows node j to receive packets successfully.

3.3 Conditions

Given the network graph $G(N, L)$ and the corresponding interference graph $G^I_c(V, E)$ of a specific channel c, we can find the conditions to determine whether a certain set of transmissions $F_c = \{l_1, l_2, \ldots, l_{|F_c|}\} \subset L$ on the same channel c is able to be active concurrently.

A necessary condition: We observe that two links incident on the same node can not operate simultaneously on the same channel. So a set of transmissions is able to be active concurrently if none of its edges is incident with each other on the same node.

A sufficient condition: Because every link in F_c interferes with each other, in order for them to be active concurrently, every receiver's SINR of all links must be greater than their designed threshold values.

In summary, we have the following definition of an admissible schedule on the same channel.

Definition 1 (Admissible Schedule): A set of concurrent transmissions $F_c \subset L$ on the same channel c is defined as an *admissible schedule* if none of links in F_c is incident with the others on the same node and their receivers' SINR values are greater than the designed thresholds.

Consequently, we have the following theorem:

Theorem 1: A set of concurrent transmissions $F_c \subset L$ on the same channel c in a given network graph is admissible if every vertex v of the corresponding interference graph satisfies:

$$\sum_{i=1}^{|F_c|} w^{v_i}_{v} < 1. \tag{3}$$

Proof Suppose that we have the set of links $\{l_i, i = 1, 2, \ldots, |F_c|\}$ transmitting simultaneously with link l. Hence, in respective interference graph, all edges incident and directed to v represents the interfering sources from all links l_i to link l. From (1), packets are received correctly at receiver of link when:

$$\frac{RSS^{tx(l)}_{rx(l)}}{W + \sum_{i=1}^{|F_c|} ISS^{tx(l_i)}_{rx(l)}} > \gamma_l$$

$$\Leftrightarrow \quad \frac{\sum_{i=1}^{|F_c|} ISS^{tx(l_i)}_{rx(l)}}{\frac{RSS^{tx(l)}_{rx(l)}}{\gamma_l} - W} < 1$$

$$\Leftrightarrow \quad \sum_{i=1}^{|F_c|} w_v^{v_i} < 1. \qquad (4)$$

□

4. Proposed Algorithm

4.1 General Desription

In this section, we propose a Priority-base STDMA Scheduling Algorithm (PASA) to construct an admissible schedule in a greedy fashion. Instead of considering the whole network graph, PASA just investigates a subgraph for which it will construct the admissible schedule. The reason to find a subgraph is to improve the link-fairness characteristic. If we consider the admissible schedules for the whole network graph, the links close to gateways have higher priority will take over the right to be scheduled first. It leads to some links at the border of system may not have a chance to transmit the data. When setting admissible schedule for a subgraph in each period, the number of high priority links has been reduced, so the border links of the network will have a chance to transmit their data with higher probability. Consequently, we decide to choose Minimum Spanning Tree (MST) as the subgraph of the network graph in our algorithm because MST is a subgraph that has all properties suitable for the purpose of our algorithm as follows.

- First, MST is a spanning subgraph that contains all links of network graph so it gives an equal chance for all links incident with all nodes to be considered in each period of the schedule.
- Second, MST of a graph defines the cheapest subset of links that keeps the graph in one connected component. So each link in a MST will have the higher priority than the others incident on the same node with it if we set the cost of each link in the sense that the higher priority, the lower cost that link is. It satisfies the condition that links with higher priority will be considered to be scheduled first.
- Finally, they can be computed quickly and easily, e.g. Kruskal's minimum spanning tree algorithm [5] can have the running time $O(|L| \log |N|)$. It's an important factor to reduce time complexity of our algorithm.

Figure 3 is an example of MST (the bold lines) constructed from a WMN. There are total 7 links that operates on channel 1 and contends to be scheduled for whole network graph while in this MST, there are just 4 links. So with the priority-based criterion, links of border nodes will have higher chance to be in a schedule.

The pseudo code of PASA is given in Algorithm 1. First, MST construction and channel assignment are done centrally by the gateway. Next, for each channel c, we now

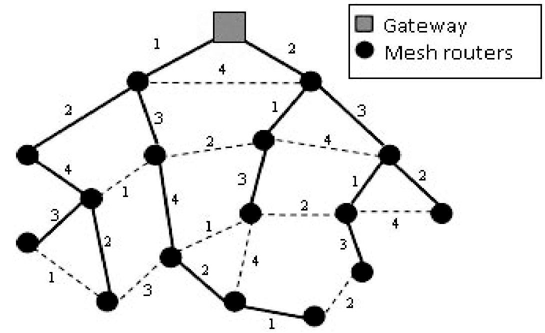


Fig. 3 A MST of a WMN with 4 orthogonal channels.

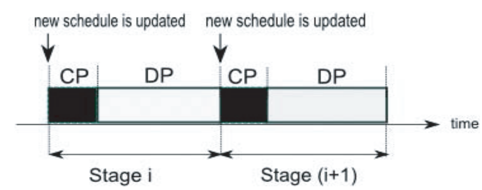


Fig. 4 Division of time into data transmission phase and control signalling phase.

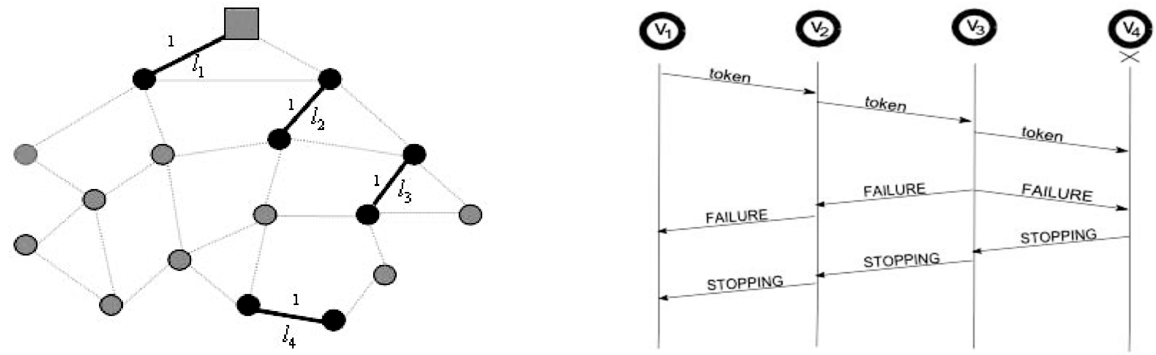
Algorithm 1 *P*riority-based STDM*A* *S*cheduling *A*lgorithm (PASA)

1: Construct a MST with K tiers
2: Do *Channel Assignment* for all channels in set C
3: **for** $c = 1$ **to** $|C|$ **do**
4: $F_c := \emptyset$; $G^I_{F_c}(V_{F_c}, E_{F_c}) := \emptyset$
5: **for** $k = 1$ **to** K **do**
6: choose the longest-queue-length link at tier k, $l^k_{Q_{max}}$
7: $V_{F_c} := V_{F_c} \cup l^k_{Q_{max}}$
8: **if** $G^I_{F_c}(V_{F_c}, E_{F_c})$ still satisfies *Theorem 1* **then**
9: $F_c := F_c \cup l^k_{Q_{max}}$
10: **end if**
11: **end for**
12: Set the duration of the schedule $T_s := \max_{l \in F_c} Q_l(t)$.
13: Each link l^k is active within its $Q^k_l(t)$ time-slots.
14: **end for**

consider a subset of links on the same channel c. At each tier, PASA chooses the longest-queue-length link as a candidate for the schedule (line 6). If its appearance does not affect any admissible links chosen before, this link will be added to the admissible schedule F_c (line 8). Otherwise, the scanning procedure finding an admissible schedule will discontinue. Finally, each link in F_c will release all of its packets in the buffer, corresponding to the number of time-slots it is active.

4.2 Detailed Operations

Each schedule generated by PASA is updated at the beginning of a stage. Due to possible collisions between control packets and data packets in shared wireless medium, the whole schedule duration can be divided into the control phase (CP) and data phase (DP) (see Fig. 4) to prevent such collisions. All nodes and links are assumed to be synchronized in the same CP and DP intervals where the synchro-



(a) 4 links of the MST operates on channel 1

(b) Vertex v_3 announces FAILURE when v_4 is added to the schedule

Fig. 5 An example of detail operations of PASA on channel 1.

nization condition can be relaxed by adding an extra guard time between CP and DP to compensate for propagation delays. All nodes (links) are given uniquely their own *IDs* by the gateway throughout the lifetime of the network. These *IDs* can be ordered according to their tier position.

In DP, links simply release all their packets according to the schedule made by PASA. Therefore, we mainly focus on how network operations are in CP. At the beginning of each stage, links send their queueing delay information (i.e. number of packets in queue according to Little's theorem) via hop-by-hop message passing from border nodes to the gateway. Any upper-tier node receives the message will add a new field to it with node (link) ID and queue size information. Therefore, along that path from border nodes to gateway, the message will include all fields of IDs and queue size. Then the gateway will use these information for MST construction (to update the link cost required by the algorithm [5]) and for channel assignment. The channel assignment problem for mesh networks is similar to the list coloring problem which is NP-complete [5]. Therefore, the gateway can use a heuristic method as follows.
At every *k*th tier of a MST subgraph, the gateway:

1. sort links by increasing delay value
2. assign higher capacity channels in set C to higher-priority links that does not conflict with their neighbors
3. if no channel is found at Step 2, assign random channels

While doing channel assignment in each schedule stage, the gateway also searches the longest-queue-length links at each tier and stores their *IDs* list orderly in a token. From now on, the interference graph will be constructed gradually by adding vertices one after another by using measurement method [18]. For that purpose, token will be passed from hand to hand by the vertices that have their *IDs* in the token's list. The gateway begins this procedure by sending token to the vertex whose *ID* is at the head of list. This is also the signal announcing this vertex to activate the control channel in broadcast mode for measurement. After checking for a while that its condition (3) is satisfied, this vertex looks into the next *ID* in the list and keeps passing the token to the corresponding vertex. This procedure continues until the last vertex at tier K finishes if there is no failure. Failure happens when any vertex added before finds that its condition (3) fails, then it will broadcast to all vertices the *FAILURE* message. The vertex currently keeping the token immediately broadcasts the *STOPPING* message including its *ID* to all of vertices in the list when receiving the *FAILURE* message. Note that the "failed" vertex is not necessarily the token-holding vertex. Finally, all vertices whose *IDs* are smaller than the *ID* in *STOPPING* message will implicitly know that they can be active in the DP.
Let us consider an example given in Fig. 5. From the MST in Fig. 3, we focus the detail operations of PASA on channel 1, and they are similarly intuitive for other orthogonal channels. In the network graph, Fig. 5(a) shows that there are 4 links working on channel 1 at different tiers of the MST. In the interference graph, the token is passed from vertex v_1 to v_4 to build the schedule greedily as shown in Fig. 5(b). When v_4 is added to the interference graph, v_3 announces that its condition (3) is not satisfied by sending *FAILURE* message. Then, v_4 which is currently holding the token broadcasts *STOPPING* message to other vertices. Thus, the final schedule on channel 1 includes v_1, v_2 and v_3.

We easily observe that apart from the MST construction and channel assignment are done centrally by gateway, our greedy method to find an admissible schedule by using interference graph can be implemented in a distributed fashion with local communications between vertices as described above. This property of PASA greatly reduces the burden on the gateway, making the algorithm practical for WMNs.

5. Algorithm Analysis

5.1 Complexity

Suppose that Kruskal's algorithm is chosen for building MST, then its running time is $O(|L| \log |N|)$. The number of tiers in the worst case is binary tree†, which has $O(\log |N|)$

†Here we exclude the case there's only one link at each tier, leading to the number of tiers is $O(|N|)$).

tiers. In channel assignment, the most computational procedure is sorting whose complexity is $O(|L|\log|L|)$ using quick sort, hence the complexity of channel assignment is $O(|L|\log|L|\log|N|)$. The operations at line 6, 7, 8 and 9 take $O(1)$ time in each tier iteration, so these operations have $O(\log|N|)$ time-complexity. Finally, the total time-complexity of PASA is $O(|L|\log|L|\log|N|)$. Furthermore, due to the distributed fashion of these operations at line 7 and 8, the number of messages exchange should also be taken into account. Since there are at most $O(\log|N|)$ tokens in the network during these operations, the total messages exchange is $O(\log|N|)$.

5.2 Throughput

In this section we consider the case of only one channel exists in WMNs without loss of generality. We denote the set of admissible link schedules by $\mathcal{M} \subset \{0,1\}^L$ and a link l in a schedule $m \in \mathcal{M}$ is defined to be active when $m_l = 1$, or $m_l = 0$ otherwise. Each schedule $m \in \mathcal{M}$ can be rendered as a state of an ergodic finite Markov process with 2^L possible states. The capacity of a link is assumed to be unit rate. We denote the stationary distribution of state m by π_m, and the transition matrix of the Markov process by $\boldsymbol{T}^{2^L\times 2^L} = [T_{m,n}], \forall m, n \in \mathcal{M}$. With the greedy maximal characteristic of our scheduling algorithm, the set $\mathcal{M}$ is a maximal schedules set. Therefore, the *maximal normalized throughput* (or service rate) of link l is:

$$\tau_l = \sum_{m\in\mathcal{M}} \pi_m m_l. \tag{5}$$

If the transition matrix is known *a priori*, the stationary distribution of all the states of Markov chain can be computed by solving the following linear equations system:

$$\sum_{m\in\mathcal{M}} \pi_m \cdot T_{m,n} = \pi_n, (\forall n \in \mathcal{M}) \tag{6}$$

$$\sum_{m\in\mathcal{M}} \pi_m = 1 \tag{7}$$

The derivation of transition matrix $\boldsymbol{T}$ is provided in Appendix. After all, the network throughput is defined:

$$C = \sum_{l\in\mathcal{L}} \tau_l, \tag{8}$$

where τ_l is computed as in (5).

5.3 Fairness

Fairness is one of the important factors when implementing a WMN, and our scheduling algorithm is carefully designed to take into account this problem. The algorithm tries to improve the chance of nodes farther away from the gateway by using MST. Furthermore, it also gives link priority based on the network-tier order in each schedule. For example, via (A· 9) in Appendix, we've had the probability that link l at tier k is active at the beginning of a new schedule:

$$p_{01}^k(l|m) = \prod_{i=1}^{k}\left(1 - \Phi\left[\frac{\log\gamma_l^i - \bar{\mu}_l^i}{\bar{\sigma}_l^i}\right]\right). \tag{9}$$

This expression exactly draws the insights of WMNs' fairness characteristics in that links closer to the gateway would have higher probabilities to be active in a new schedule, then it should be allocated more resources due to their higher loads of relaying traffics. If fairness issues are not considered seriously when designing WMNs, the queue stability is broken down, leading to buffer overflow and congestions at these links.

6. Performance Evaluation

In this section, we compare the performance of our proposed scheme with the target algorithm, GreedyPhysical [4], and IEEE802.11 media access scheme which employs CSMA/CA. We first describe the setting details under which simulations are carried out in Sect. 6.1. The main results proving the improvements of our algorithm on throughput and fairness are presented next in Sect. 6.3 and Sect. 6.5.

6.1 Simulation Environment

We conducted extensive simulations using *ns*-2 simulator for different binary-tree network topologies by varying the number of nodes. Figure 6 shows a representative binary-tree topology with 3 tiers of 15 nodes. In all topologies, we chose the root node as the gateway. Simulations were carried out for a rectangular region of 670 × 670 meters. The routing protocol used in all simulations was Ad-hoc On-demand Distance Vector (AODV). The interface queues at each node used a Droptail policy and the queue length limit was set at 200 packets. Simulations were run at a duration of 200 seconds. Other parameter settings for the simulations are given in Table 1 where the loss factor of wireless channel is set to 1 (i.e. no loss). We only consider the upstream traffic in all scenarios, where every node except the gateway was the source of one flow and the gateway is the only sink for all flows. Thus for N nodes we have $N-1$ active flows. Each constant bit rate (CBR) traffic source used UDP as the transport protocol and the packet size was set at 1000 bytes. In all simulation scenarios, all CBR applications are set to start at time 1.0 until the simulation stops. All nodes used

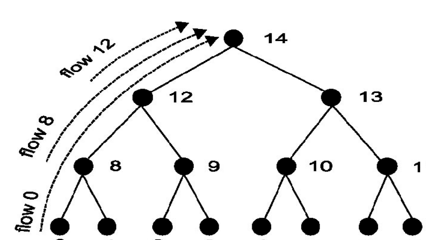


Fig. 6 A 3-tier binary-tree topology with 14 mesh routers as 14 traffic sources, 1 gateway as a sink.

Table 1 Simulation parameters.

Carrier-Sensing Threshold	−82 dBm
Power	0.001 Watt
Frequency	5.18 GHz
Noise floor	−96 dBm
Loss factor	1.0
Propagation Model	TwoRayGround

one 802.11 radio for CP and many radios for DP where all of them tuned to orthogonal channels. The beacon-message exchanges between mesh routers and the gateway were set to 2 μs. The MST construction at gateway took 10 μs for algorithm computation. The total time required to collect all queue information at gateway is set to 50 μs. Two Modulation and Coding Schemes (MCS) used in most scenarios are BPSK 1/2 and QPSK 1/2, which yield a raw physical bandwidth of about 6 Mbps and 12 Mbps correspondingly per 20 MHz channel. The basic communication rate for CP was fixed to 6 Mbps while the data transmission rates were varied according to the raw physical bandwidth of different MCSs.

6.2 Performance Metrics

Several metrics are used to assess the algorithm performance. *End-to-end throughput* and *MAC throughput* are considered as the main indices for evaluating throughput performance. While end-to-end throughput of a traffic flow is the number of bits received by the gateway per second without MAC overhead, the MAC throughput of a node is the number of bits received at MAC layer of that node per second, regardless of the traffic flow to which the received data belong.

Relating to fairness, the *transmission probability* of a mesh router is the probability accessing channel of that node for whole simulation duration. Also, the *fairness index* is considered in evaluating the even bandwidth partitioning among traffic flows. Fairness index is defined as the ratio $(\sum_{i=1}^{n} x_i)^2/(n\sum_{i=1}^{n} x_i^2)$, where n is the number of traffic flows, and x_i is the throughput of the i-th traffic flow.

6.3 Throughput Improvement Effect

We first consider the single-channel case, where the sum of end-to-end throughput and MAC throughput are shown in Fig. 7 and Fig. 8. Different MCSs are employed for comparison, namely BPSK 1/2, QPSK 1/2.

In case of end-to-end throughput (see Fig. 7), we can see as the number of nodes increases, the end-to-end throughput decreases for any MCSs. It can be explained that an increasing fraction of channel capacity is utilized to relay packets at intermediate nodes. For instance, with three nodes, all PASA, GreedyPhysical and IEEE802.11 achieve the same throughput which is approximate to the raw bandwidth because there is no advantage of spatial-reuse in this case, where two flows on two links alternatively share the wireless channel. However, with seven nodes corresponding

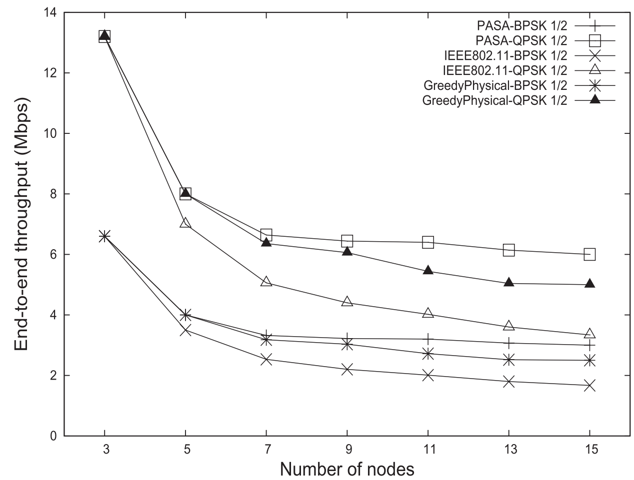


Fig. 7 End-to-end throughput of all traffic flows.

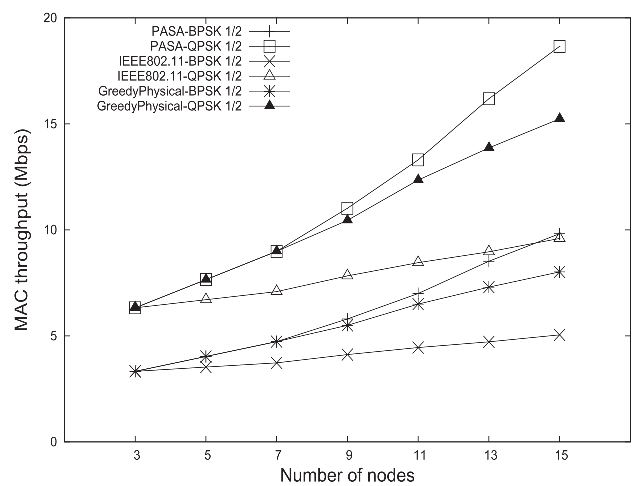


Fig. 8 MAC throughput of all nodes.

to 2 tiers, while PASA and GreedyPhysical whose throughput are nearly equal can achieve approximately 2/3 of the raw bandwidth, of which IEEE802.11 only achieve 1/2. This phenomenon is due to the fact that with IEEE802.11, most wireless resources are employed to transmit data to upstream neighbors in a random access fashion without any priority concern, which creates many bottlenecks along the network and block those flows that are farther from the gateway, hence severely degrades their throughput performance. In contrast, with the load-awareness STDMA characteristics, PASA and GreedyPhysical have overcome this negative effect. When the number of node continues increasing, we observe that the throughput of GreedyPhysical degrades more severely than that of PASA. For example, when the number of nodes is more than 11 corresponding to the number of tiers of WMNs is 3, while IEEE802.11's throughput is still 1/2 of PASA, GreedyPhysical's throughput is only about 2/3 of PASA's. Now we can see that PASA outperforms GreedyPhysical when the size of WMNs increases since the priority-based fairness comes into effect. When the traffic is heavily loaded in the system, instead of treating all links the same priority as GreedyPhysical leading to

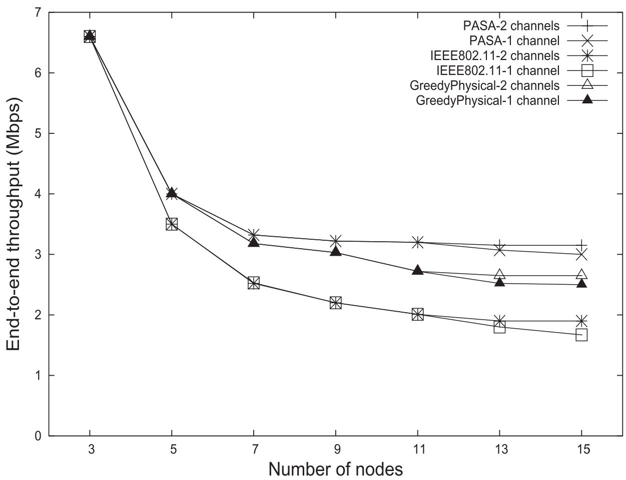


Fig. 9 End-to-end throughput of all traffic flows.

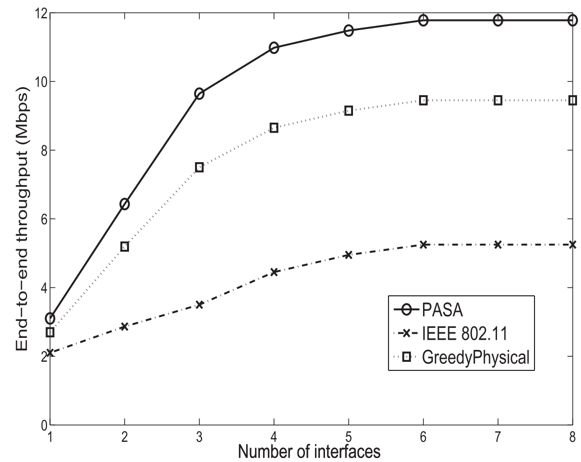


Fig. 10 Aggregated end-to-end throughput with 8 channels and different number of interfaces.

the bottlenecks at some high loaded links, PASA gives more opportunistic schedule to links closer to the gateways to mitigate the bottleneck and hence improves the throughput. We also observe that the end-to-end throughput is proportional to the MCS efficiency; for three of PASA, GreedyPhysical and IEEE 802.11 access scheme, the BPSK 1/2 throughput is about half of QPSK 1/2 throughput.

In case of MAC throughput, we easily see that the MAC throughput increases with the number of nodes. It can be explained that when the number of nodes increases, the amount of packets transmitted in the network also rise up. When the number of nodes is less than 5 where there is not much spatial-reuse utilization, the MAC throughput performance of PASA, GreedyPhysical and IEEE 802.11 are almost the same. Apart from that, the MAC throughput of PASA and GreedPhysical is significantly higher than that of IEEE 802.11 since with the capability of spatial-reuse. So PASA and GreedyPhysical are better in exploiting the channel utilization. Again, when the number of nodes is more than 11, MAC throughput of PASA is higher than GreedyPhysical due to the priority-base link-fairness provision of PASA, which is efficient in alleviating the bottlenecks in high traffic demand WMNs. Also, as can be seen the MCS efficiency of MAC throughput is similar to that of end-to-end throughput for all PASA, GreedyPhysical and IEEE 802.11 access scheme.

Next, we consider the performance of multi-channel WMN by setting each node two orthogonal channels. Now we have one radio interface with two orthogonal channels in DP. As can be seen in Fig. 9, adding one more channel did not double the end-to-end throughput as we expected. This is due to the fact that each node has only single radio, thus the data transmission of links in the same schedule on different orthogonal channels can happen between disjoint pairs.

6.4 Multi-Radio Multi-Channel Effect

While in previous section we focus on the throughput improvement of PASA with different MCS but with only one single radio for DP, in this section we will generalize the effect of multi-radio multi-channel in WMN by setting many interfaces for DP with many orthogonal channels. The MCS used in this section is BPSK 1/2. First we fix $|C| = 8$, and $I \in \{1, \ldots, |C|\}$. As we can see from Fig. 10, the aggregated end-to-end throughput (of 14 flows) increases as the number of interfaces increases. Moreover, the additional throughput gained by adding a new card decreases as the number of cards increases. For instance, with $|C| = 8$, only 5 radios are enough to achieve the maximal throughput (i.e., maximal throughput is achieved when $|C| = |I|$) for three schemes. These maximal throughputs are 11.8, 9.5 and 5.25 Mbps respectively for PASA, GreedyPhysical and 802.11. This observation is consistent with the result in [11], which showed the relation between network throughput and ratio of channels to interfaces. Again we clearly see that performance of PASA and GreedyPhysical far better than IEEE 802.11 in this multi-radio multi-channel setting.

Figure 11 shows the ratio gain of the aggregated end-to-end throughput of 3 MAC schemes as a function of the number of channels and interfaces. The ratio gain is with respect to the case $|I| = 1, |C| = 1$ (c.f. Sect. 6.3). As can be seen from the figure, due to spatial reuse, PASA and GreedyPhysical have impressing ratio gain 2 (i.e. double throughput) approximately compared with only 1.35 of IEEE 802.11 in case of $|I| = 2$. Similarly, with other cases of $|I| = 3$ and $|I| = 4$, the ratio gains of PASA are 3 and 3.5, while GreedyPhysical are 2.8 and 3.2, and IEEE 802.11's gains are only 1.7 and 2.

6.5 Fairness Enhancement Effect

We first analyze the results in Fig. 12, which shows the node transmission probabilities in the WMN with 15 nodes. It is not surprising that by utilizing IEEE 802.11, nodes far away from the gateway cannot be guaranteed to have higher channel-access probabilities as employing PASA, which supports our expectation. It is clear that the priority-base

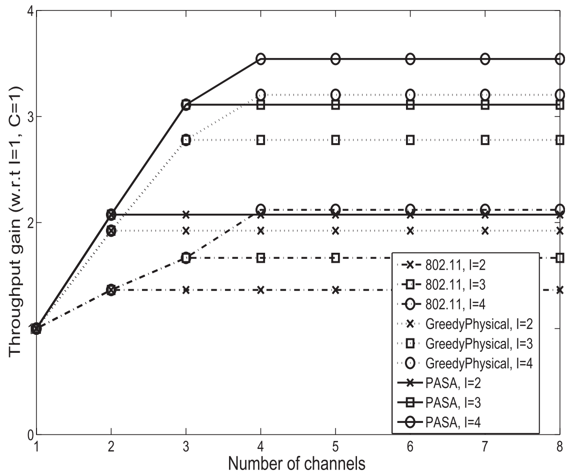


Fig. 11 Ratio gain factor in terms of aggregated end-to-end throughput with respect to case single radio, single channel.

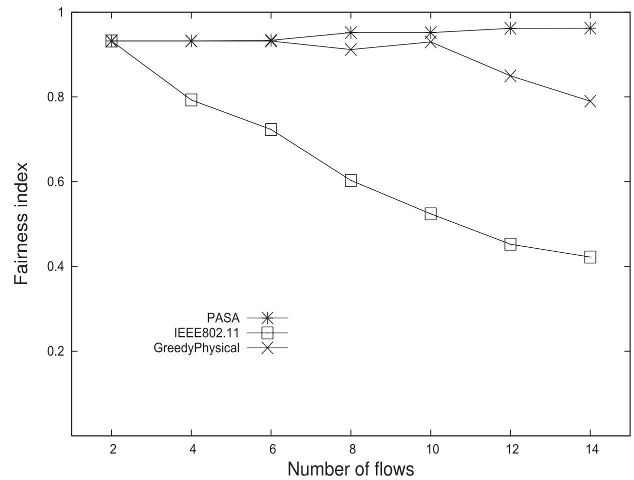


Fig. 13 Fairness index of end-to-end throughput.

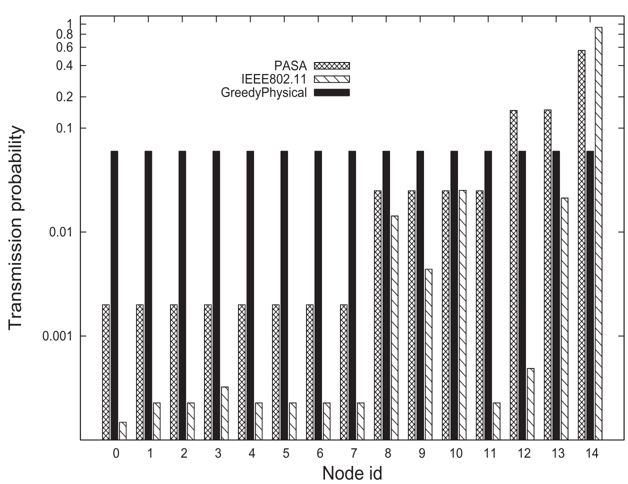


Fig. 12 Nodes' transmission probabilities.

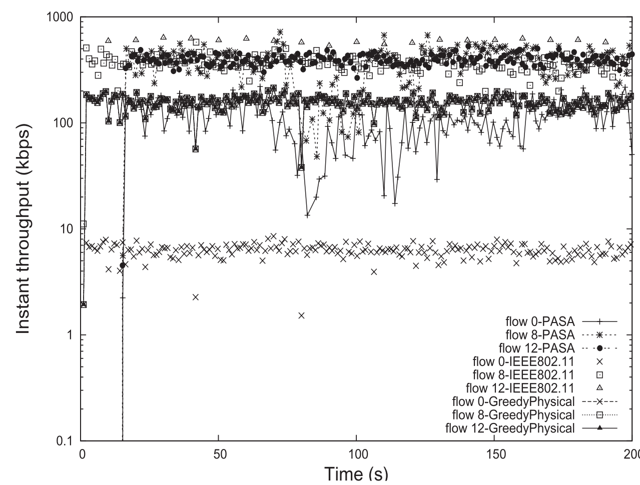


Fig. 14 Instant throughputs of flow 0, 8 and 12.

essence of PASA provides equal transmission probabilities between nodes on the same tiers. For instance, nodes with IDs from 0 to 7 in tier 3, from 8 to 11 in tier 2, and from 12 to 13 in tier 1 have nearly the same transmission probability correspondingly to 0.003, 0.05 and 0.18. In contrast, nodes employing IEEE 802.11 MAC have their transmission probabilities in a random order due to its random access characteristic. For instance, even though the node with ID 11 is in the same tier as nodes with IDs 8, 9 and 10, its transmission probability is very low to compare with those three nodes. In case of GreedyPhysical, all nodes have equal accessing probability 0.07 approximately, which reflects exactly the spirit of this algorithm.

We next consider about the fairness index presented in Fig. 13. The results again strongly support the fairness efficiency of PASA not only in link fairness but also in flow fairness. As can be seen PASA is the algorithm that achieve nearly perfect fairness among all traffic flows regardless the number of nodes. On the other hand, IEEE 802.11 is only fair when the number of flows is small while GreedyPhysical's fairness is degraded when the number of flows increases over a threshold. For instance, when there are only three nodes (i.e., two flows) in the system, the fairness index of PASA, GreedyPhysical and IEEE 802.11 are the same at 0.95. When the number of nodes increases, the fairness index of IEEE 802.11 sharply decreases, which is already mentioned above by the starvation of nodes (links) that are far away from the gateway while that of GreedyPhysical slightly decreases when number of flows is over 10. In order to illustrate more about fairness index, Fig. 14 shows the instant throughput of three flows whose sources originate from nodes 0, 8 and 12. These flows represent for traffic from different tiers in a WMN as in Fig. 6. While these three flows in case of GreedyPhysical have nearly the same throughput, flow 0 whose source is farthest from the gateway has only 1/100 of the instant throughput of flow 12 whose source is closest to the gateway in case of IEEE 802.11. And finally, three flows of PASA are in the order corresponding to their source's position.

7. Conclusion

In this paper, we propose a greedy scheduling algorithm whose goals are enhancing throughput and fairness. Realizing that CSMA/CA or TDMA approaches is not a suitable medium access scheme for WMNs, we utilize STDMA scheme for our scheduling algorithm to exploit its spatial-reuse characteristic, which can guarantee that links can be given more chances to access the medium, and hence increase the throughput. Besides the throughput consideration, fairness is also taken into account in that links closer to the gateway have more opportunity to be active because of having heavy loads. We also provide the analysis for our scheduling scheme and extensive simulation results are also given as a evidence to validate our analysis.

References

[1] F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," Elsevier Journal of Computer Networks, vol.47, no.4, pp.445–487, 2005.
[2] M. Alicherry, R. Bathia, and L. Li, "Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks," ACM MOBICOM, pp.58–72, 2005.
[3] P. Bjorklund, P. VWmd, and D. Yuan, "Resource optimization of spatial TDMA in ad hoc radio networks: A column generation approach," IEEE INFOCOM, pp.818–824, 2003.
[4] G. Brar, D. Blough, and P. Santi, "Computationally efficient scheduling with the physical interference model for throughput improvement in wireless mesh networks," ACM MOBICOM, pp.48–62, 2006.
[5] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, Introduction to Algorithms, 2nd ed., MIT Press, 2001.
[6] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," ACM MOBICOM, pp.114–128, 2004.
[7] L.F. Fenton, "The sum of lognormal probability distributions in scatter transmission systems," IRE Trans. Commun. Syst., CS-8, 1960.
[8] J. Gronkvist, J. Nilsson, and D. Yuan, "Throughput of optimal spatial reuse TDMA for wireless ad-hoc networks," IEEE VTC, pp.2156–2160, 2004.
[9] A. Hamed, M. Rad, and V. Wong, "Joint channel allocation, interface assignment and MAC design for multi-channel wireless mesh networks," IEEE INFOCOM, 2007.
[10] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," ACM MOBICOM, pp.66–80, 2003.
[11] P. Kyasanur and N. Vaidya, "Capacity of multi-channel wireless networks: Impact of number of channels and interfaces," ACM MOBICOM, pp.43–57, 2005.
[12] J. Lee, W. Liao, and M. Chen, "An incentive-based fairness mechanism for multi-hop wireless backhaul networks with selfish nodes," IEEE Trans. Wireless Commun., vol.7, no.2, pp.697–704, Feb. 2008.
[13] T. Liu and W. Liao, "Location-dependent throughput and delay in wireless mesh networks," IEEE Trans. Veh. Technol., vol.57, no.2, pp.1188–1198, March 2008.
[14] T. Moscibroda and R. Wattenhofer, "The complexity of connectivity in wireless networks," IEEE INFOCOM, 2006.
[15] R. Nelson and L. Kleinrock, "Spatial-TDMA: A collison-free multi-hop channel access protocol," IEEE Trans. Commun., vol.33, no.9, pp.934–944, Sept. 1985.
[16] K.S. Kim, S. Uno, and M.W. Kim, "Adaptive QoS mechanism for wireless mobile network," J. Computing Science and Engineering, vol.4, no.2, pp.153–172, June 2010.
[17] A. Raniwala and T. Chiueh, "Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh networks," IEEE INFOCOM, pp.2223–2234, 2005.
[18] C. Reis, R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan, "Measurement-based models of delivery and interference," ACM SIGCOMM, 2006.
[19] N. Salem and J. Hubaux, "A fair scheduling for wireless mesh networks," IEEE WIMESH, 2005.

Appendix: Computing Transition Matrix *T*

Under our proposed scheduling scheme, the state transitions of every link are dependent on each other. In general, however, a typical state transition of a link across timelots involves four possible cases as in Fig. A· 1 corresponding to computing four transition probabilities in every given state m: (a) keeping inactive, $p_{00}(l|m)$; (b) being active, $p_{01}(l|m)$; (c) returning inactive $p_{10}(l|m)$ and (d) keeping active, $p_{11}(l|m)$. Then, each entry of transition matrix T can be derived as follows:

$$(\forall m, n \in \mathcal{M}), \quad T_{m,n} = \prod_{\{l|m_l=0,n_l=0\}} p_{00}(l|m) \cdot \prod_{\{l|m_l=0,n_l=1\}} p_{01}(l|m) \cdot \prod_{\{l|m_l=1,n_l=0\}} p_{10}(l|m) \cdot \prod_{\{l|m_l=1,n_l=1\}} p_{11}(l|m). \qquad \text{(A· 1)}$$

The tasks now are to compute in turn each of four transition probabilities. First, in order to deal with $p_{01}(l|m)$, we recall that in our scheduling scheme, at each step one link of tier k is only considered to be in a schedule when all links of upper tiers are already admissible. Therefore, we represent $p_{01}(l|m)$ by new notation $p^k_{01}(l|m)$ to indicate this link is in kth tier and we denote by ε^k the event $SINR^k_l > \gamma^k_l$. Then we have:

$$\begin{aligned} p^k_{01}(l|m) &= Pr\{\varepsilon^k \wedge \varepsilon^{k-1} \wedge \ldots \wedge \varepsilon^1\} \\ &= Pr\{\varepsilon^1\} \cdot Pr\{\varepsilon^2|\varepsilon^1\} \ldots Pr\{\varepsilon^k|\varepsilon^{k-1} \ldots \varepsilon^1\} \\ &= \prod_{i=1}^{k} Pr\{SINR^i_l > \gamma^i_l\}. \end{aligned} \qquad \text{(A· 2)}$$

Now we focus on computing $Pr\{SINR_l > \gamma_l\}$ for an arbitrary link l. We denote the RSS and ISS at the receiver of

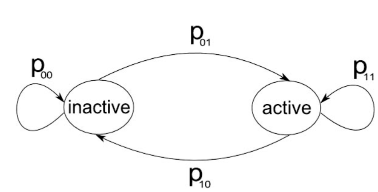


Fig. A· 1 State transitions of a link.

link l by random variables R_l and $\sum_{\{l' \in m | m_{l'}=1\}} I_l^{l'}$, and the total interference at link l in a given state m by $Y_{l|m}$, where:

$$Y_{l|m} = W_l + \sum_{\{l' \in m | m_{l'}=1\}} I_l^{l'} \tag{A$\cdot$3}$$

By assuming all $I_l^{l'}, \forall l' \in m$ are independent, we can estimate this sum by using the method of [7], where each term is modeled as a log-normal random variable then the sum itself is also approximated as a log-normal random variable. Suppose that the sum of noise and interference at link l can be represented by a log-normal random variable e^Z with $Z \sim N(\mu, \sigma^2)$, then by equating the first two moments of e^Z and $Y_{l|m}$, we have:

$$E[e^Z] = E[Y_{l|m}] \Leftrightarrow e^{\mu+\sigma^2/2} = E[Y_{l|m}]$$
$$E[e^{2Z}] = E[Y_{l|m}^2] \Leftrightarrow e^{2\mu+2\sigma^2} = Var[Y_{l|m}] + (E[Y_{l|m})^2$$

where $E[Y_{l|m}] = W_l + \sum_{\{l' \in m | m_{l'}=1\}} E[I_l^{l'}]$, and $Var[Y_{l|m}] = \sum_{\{l' \in m | m_{l'}=1\}} Var[I_l^{l'}]$. After all, we have:

$$\mu = 2 \log E[Y_{l|m}] - \frac{1}{2} \log E[Y_{l|m}^2] \tag{A$\cdot$4}$$
$$\sigma^2 = \log E[Y_{l|m}^2] - 2 \log E[Y_{l|m}]. \tag{A$\cdot$5}$$

After approximating the total interference $Y_{l|m}$ by a log-normal random variable e^Z with $Z \sim N(\mu, \sigma^2)$ as above, we continue approximate the ratio $\frac{S_l}{Y_{l|m}}$ also by a log-normal random variable $e^{\bar{Z}}$, where $\bar{Z} \sim N(\bar{\mu}_l, \bar{\sigma}_l^2)$ because S_l is independent with e^Z and the ratio of independent log-normal random variables also has log-normal distribution. Then we have:

$$\bar{\mu}_l = E[\log S_l] - \mu \tag{A$\cdot$6}$$
$$\bar{\sigma}_l^2 = Var[\log S_l] + \sigma^2. \tag{A$\cdot$7}$$

Then, we have the result:

$$Pr\{SINR_l > \gamma_l\} = Pr\left\{\frac{S_l}{Y_{l|m}} > \gamma_l\right\}$$
$$\approx Pr\{e^{\bar{Z}} > \gamma_l\} = 1 - \Phi\left[\frac{\log \gamma_l - \bar{\mu}_l}{\bar{\sigma}_l}\right], \tag{A$\cdot$8}$$

where $\Phi[u] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-x^2/2} dx$ represents for the CDF of standard normal distribution.

Finally, from (A$\cdot$2), the final approximation for $p_{01}^k(l|m)$ is:

$$p_{01}^k(l|m) \approx \prod_{i=1}^{k} \left(1 - \Phi\left[\frac{\log \gamma_l^i - \bar{\mu}_l^i}{\bar{\sigma}_l^i}\right]\right) \tag{A$\cdot$9}$$

Then, for the probability of keeping in inactive mode of link l, we have:

$$p_{00}^k(l|m) = 1 - p_{01}^k(l|m)$$
$$= 1 - \prod_{i=1}^{k} \left(1 - \Phi\left[\frac{\log \gamma_l^i - \bar{\mu}_l^i}{\bar{\sigma}_l^i}\right]\right) \tag{A$\cdot$10}$$

Relating to computing $p_{11}^k(l|m)$, we assume that the transmission time of a link follows the exponential distribution where the mean transmission time is denoted by t_Q (the value $\lfloor t_Q \rfloor$ also can be rendered as the mean number of packets in queue of l at the beginning of a schedule), thus we have:

$$p_{11}^k(l|m) = Pr\{X > t_{slot}\} = e^{-\frac{t_{slot}}{t_Q}}, \tag{A$\cdot$11}$$

where X is the transmission time random variable and t_{slot} is the fix duration of one time-slot.

Clearly, we have:

$$p_{10}^k(l|m) = 1 - p_{11}^k(l|m) = 1 - e^{-\frac{t_{slot}}{t_Q}}. \tag{A$\cdot$12}$$

From (A$\cdot$10), (A$\cdot$9), (A$\cdot$12), (A$\cdot$11) and (A$\cdot$1), we completely derive the transition matrix $\boldsymbol{T}$.

Nguyen H. Tran received his B.E. degree in electrical engineering at University of Technology, Hochiminh city, Vietnam in 2005. Since 2006, he has been working towards the Ph.D. degree at Department of Computer Engineering, Kyung Hee University, South Korea. His research interests include distributed and randomized algorithms, cross-layer design, wireless scheduling design and stochastic network optimization.

Choong Seon Hong received his B.S. and M.S. degrees in electronic engineering from Kyung Hee University, Seoul, Korea, in 1983, 1985, respectively. In 1988 he joined KT, where he worked on Broadband Networks as a member of the technical staff. From September 1993, he joined Keio University, Japan. He received the Ph.D. degree at Keio University in March 1997. He had worked for the Telecommunications Network Lab, KT as a senior member of technical staff and as a director of the networking research team until August 1999. Since September 1999, he has been working as a professor of the School of Electronics and Information, Kyung Hee University. He has served as a Program Committee Member and an Organizing Committee Member for International conferences such as NOMS, IM, APNOMS, E2EMON, CCNC, ADSN, ICPP, DIM, WISA, BcN and TINA. His research interests include ad hoc networks, network security and network management. He is a member of IEEE, IPSJ, KIPS, KICS and KIISE.

Sungwon Lee received the Ph.D. degree from Kyung Hee University, Korea. He is a professor of the Computer Engineering Departments at Kyung Hee University, Korea. Dr. Lee was a senior engineer of Telecommunications and Networks Division at Samsung Electronics Inc. from 1999 to 2008. He is a editor of the Journal of Korean Institute of Information Scientists and Engineers : Computing Practices and Letters.