End-to-End Performance and Fairness in Multihop Wireless Backhaul Networks *

Violeta Gambiroza, Bahareh Sadeghi, and Edward W. Knightly
Department of Electrical and Computer Engineering
Rice University
Houston, TX 77005
{violeta,bahar,knightly}@rice.edu

ABSTRACT

Wireless IEEE 802.11 networks in residences, small businesses, and public "hot spots" typically encounter the *wireline* access link (DSL, cable modem, T1, etc.) as the slowest and most expensive part of the end-to-end path. Consequently, network architectures have been proposed that employ multiple wireless hops in route to and from the wired Internet. Unfortunately, use of current media access and transport protocols for such systems can result in severe unfairness and even starvation for flows that are an increasing number of hops away from a wired Internet entry point. Our objective is to study fairness and end-to-end performance in multihop wireless backhaul networks via the following methodology. First, we develop a formal reference model that characterizes objectives such as removing spatial bias (i.e., providing performance that is independent of the number of wireless hops to a wire) and maximizing spatial reuse. Second, we perform an extensive set of simulation experiments to quantify the impact of the key performance factors towards achieving these goals. For example, we study the roles of the MAC protocol, end-to-end congestion control, antenna technology, and traffic types. Next, we develop and study a distributed layer 2 fairness algorithm which targets to achieve the fairness of the reference model without modification to TCP. Finally, we study the critical relationship between fairness and aggregate throughput and in particular study the fairness-constrained system capacity of multihop wireless backhaul networks.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless communication*

General Terms

Design, Performance

*This research is supported by NSF ITR Grants ANI-0331620 and ANI-0325971, the Texas Advanced Technology Program, a Sloan Fellowship, and by Intel Corporation.

MobiCom'04, Sept. 26-Oct. 1, 2004, Philadelphia, Pennsylvania, USA.

Keywords

CSMA, CSMA/CA, Fairness, Wireless Backhaul Networks

1. INTRODUCTION

Today, commercial wireless LANs can achieve throughputs of 54 Mb/sec and beyond. Yet, for residences and public places, throughputs remain dismally slower due to slow *wired* backhaul connections in the hundreds of kb/sec range as provided by DSL or cable modems. While achieving higher-speed wireline backhaul for residences and "hot spots" is technically feasible, it is unfortunately not *economically* feasible to match the capacity of the backhaul link to that of the wireless LAN. Consequently, a number of research and commercial efforts are developing *wireless* backhaul networks that forego costly wired infrastructure via wirelessly multi-hopping to a high-speed and low-cost wired Internet entry point such as a metropolitan network operations center or a university [5, 17]. Figure 1 illustrates such an architecture in which traffic is forwarded over multiple wireless "Transit Access Points" (TAPs) in route to or from the wired Internet.

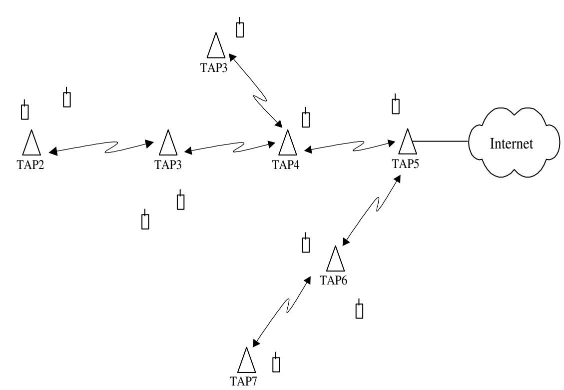


Figure 1: Illustration of Multihop Wireless Backhaul

Unfortunately, current protocols are severely inadequate in achieving the design goals of multihop wireless backhaul networks. In particular, existing protocols result in severe unfairness, poor performance, and in some cases, starvation, for users located more than one hop away from the wired entry point.

The contributions of this paper are as follows. First, we provide a formal reference model that characterizes the idealized fairness and throughput objectives for multihop wireless backhaul networks. The model differs from classical max-min fairness [1] as well as proportional fairness as targeted by TCP [18] in that it (i) does not penalize multihop flows vs. spatial and RTT bias, (ii) achieves fairness at the TAP-aggregate granularity vs. flow granularity,[1] (iii) is temporal fair vs. throughput fair as is essential for multi-rate wireless links [12, 28, 29], and (iv) maximizes spatial reuse subject to the first three constraints.

Next, we perform an extensive set of simulation experiments to characterize the relative performance impact of the key factors that influence fairness and capacity including: (i) the media access protocol and use of CSMA with and without Collision Avoidance, (ii) the use of sector antennas to increase spatial reuse and system capacity, (iii) the use of TCP to achieve fairness and capacity objectives, (iv) multi-rate channels and their impact on fairness.

Next, we develop a simple layer-2 Inter-TAP Fairness Algorithm (IFA) which seeks to achieve the reference model's goals via a distributed algorithm operating purely within the multihop wireless network as compared to TCP's end-to-end approach. The key mechanism of IFA is a local fairness computation at each TAP that is forwarded upstream such that a flow is throttled at its ingress TAP to its network-wide fair rate. While complete development of such an algorithm that would consider all implementation details is beyond the scope of this work, we utilize IFA to experimentally study this class of solutions. In addition to system factors (i)-(iv) above, we consider the impact of the joint use of TCP and IFA on fairness and capacity. Moreover, we show that IFA provides *inter-node* performance isolation: a TCP or UDP flow from any TAP is ensured its fair share even in the presence of non-responsive traffic originating from other TAPs.

Finally, we utilize the fairness reference model to explore fairness-constrained system capacity. In particular, previous studies such as [15] consider network capacity *without* fairness constraints, which can result in significant unfairness were a protocol to realize such a capacity: for example, in certain scenarios a capacity-maximizing strategy would give all bandwidth to one-hop flows and starve multi-hop flows. Thus, we explicitly characterize the link between fairness and capacity with a particular focus on multihop wireless backhaul networks (i.e., we do not consider general ad hoc networks).

Our main experimental findings are as follows. First, we find that scenarios such as in Figure 1 result in "hidden terminals" and "information asymmetry" which results in near starvation of upstream flows for any combination of UDP, TCP, CSMA, and CSMA/CA. On the other hand, the use of sector antennas mitigates this problem yet still results in throughputs as low as 24% of the targeted value of the reference model due to inefficiencies and significant spatial bias introduced by TCP. In contrast, simulations with the Inter-TAP Fairness Algorithm show that a multi-hop layer 2 algorithm can achieve near-perfect fairness and 76% and 71% of the total available capacity for CSMA and CSMA/CA respectively. As 100% throughput is not feasible due to collisions and retransmissions, such throughputs are quite close to the maximum achievable under such MACs [3]. Moreover, throttling flows at their ingress points to their system-wide fair rate has the side effect of mitigating the effects of hidden terminals and information asymmetry, without any modification to CSMA nor CSMA/CA. Finally, with TCP and IFA, flows obtain throughputs of 59% to 75% of that targeted by the reference model, without requiring any changes to TCP Sack.

The remainder of this paper is organized as follows. Next, in Section 2 we present the fairness reference model. We turn to simulation experiments with current protocols establishing baseline simulations in Section 3 and considering multiple performance factors in Section 4. In Section 5, we develop IFA and evaluate its performance. In Section 6, we explore the relationship between system capacity and fairness under multiple fairness reference models. Finally, in Section 7 we review related work and in Section 8 we conclude.

[1] That is, the targeted service granularity is per-customer (residence, hot spot, etc.) vs. per-TCP-micro-flow. Within a customer's ingress aggregate flow, sub-flows are also treated fairly.

2. PERFORMANCE AND FAIRNESS OBJECTIVES

In this section, we present background on fairness and devise a formal definition of the fairness objectives for multihop wireless backhaul networks.

2.1 Background on Fairness

A fairness reference model provides a formal idealized objective that can be used as a target and benchmark for protocol design and as a tool for studying alternatives for a network's fairness and performance objectives.

For a single *wired* node, the fairness objective is immediate and is defined by max-min fairness [1] and realized by fair queueing [26]. Yet for a *wireless* network, even in a simple case with a single access point, the fairness objective must consider the *resource* which is to be fairly allocated. If the resource is *throughput*, then IEEE 802.11 performance degrades considerably as all flows match their throughput to that of the flow with the lowest quality channel (see [12]). On the other hand, if the resource is *time* (see [29]), then all users are assured an equal time share of channel access, so that users with high quality channels can obtain throughput gains independent of the channel qualities of others. Likewise, users with poor channels are also guaranteed their fair time share. In this way, time-share fairness provides the desirable "performance isolation" property and avoids the "performance anomaly" of throughput fairness.

For *multihop wireline* networks, there are multiple possible fairness objectives, including extending max-min fairness to multiple resources [1] as targeted in the ATM fairness literature and proportional fairness [18, 23] as targeted by TCP, i.e., TCP achieves throughput inversely proportional to round-trip-time thereby penalizing longer-distance higher-path-length flows. In *multihop wireless* networks, defining the fairness objective must also address contention neighborhoods and variable rate channels.

Yet, despite significant progress in fairness in multihop wireline and wireless networks, no existing reference model captures the design objectives of multihop wireless backhaul networks. Thus, we formally define such a system's fairness and throughput objectives as follows.

2.2 TAP Fairness Reference Model

The reference model for fairness in multihop wireless backhaul networks has the following four objectives. First, the targeted granularity of fairness is a TAP-aggregated flow. In particular, each TAP corresponds to a single residence, small business, or hot spot, and this TAP's traffic should be treated as a single aggregate, independent of the number of TCP micro-flows or mobile devices supported by the TAPs. While fairness can be weighted among TAPs according to pricing or other system-wide objectives, the basic fairness granularity is a per-TAP aggregate.

Second, maximal spatial reuse must be ensured subject to the first constraint. That is, network resources can be reclaimed by TAP-aggregated flows when they are unused either due to lack of demand or in cases of sufficient demand in which flows are bottlenecked elsewhere.

Third, spatial bias must be eliminated to ensure that nodes one hop away from a wired entry point do not receive a disproportionately greater share of resources than nodes multiple hops away.

This property is essential for deployability of multihop wireless backhaul architectures to ensure that users do not suffer a performance penalty for not having a wireline Internet connection.

Fourth, *time* rather than throughput should be considered as the basic network resource that needs to be fairly shared.

In the Appendix, we present a precise mathematical formulation for the above design objectives. In addition to the above considerations, the reference model considers the shared-medium aspect of wireless networks, namely, that collections of nodes are divided into contention neighborhoods in which, for example, received traffic contends with outgoing traffic (unlike wireline networks). Thus, this fairness reference model differs from classical objectives in the fairness granularity (TAP aggregated vs. per-node or per-flow), the basic resource considered (time vs. throughput), spatial properties (no spatial bias and maximum spatial reuse), and the medium (multirate shared wireless channels vs. wired links).

Finally, note that there is a critical link between fairness and total system throughput (i.e., the sum of all flow throughputs). This relationship is established in Section 6 and further discussion of the fairness literature is presented in Section 7.

3. UDP BASELINE SCENARIO

Our experimental objective is to study the key performance factors of end-to-end performance and fairness via consideration of multiple fairness algorithms [uncontrolled (continuously backlogged UDP), TCP-SACK, and IFA], media access protocols [CSMA and CSMA/CA], channel models [constant rate, Ricean], antenna technologies [omni directional, sector], traffic types [continuously backlogged, variable rate], as well as multiple topologies and flow scenarios. We present results from a fractional factorial experimental design that considers most combinations (and hence interactions) of the above factors with average results of multiple 50 second ns-2 simulation runs reported.

3.1 Scenario

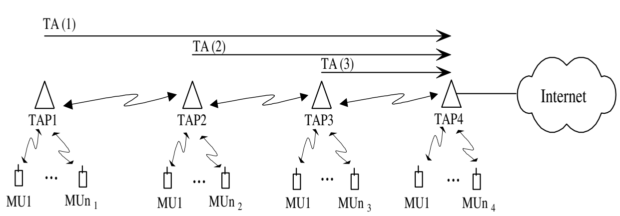


Figure 2: Parking Lot Scenario

In this section we establish a baseline scenario for experimental performance analysis that consists of the first factor in each set above and with a "parking lot" topology and flow scenario as depicted in Figure 2.[2] The name of the scenario is intended to convey a situation analogous to many cars simultaneously attempting to leave a parking lot and the resulting congestion and lack of fairness that ensues. This scenario has 4 TAPs, each serving a different number of wireless devices which we refer to as Mobile Units (MUs), and all traffic is destined to the wired Internet. Note that this scenario represents a single branch in the access tree depicted in Figure 1. If different branches are placed on orthogonal frequencies or are sufficiently spatially separated, then the results obtained for a single branch represent system-wide behavior. On the other hand, if different branches are within radio range (an issue beyond the scope of this study), then throughputs will be correspondingly lower due to increased contention and interference. In any case, we will consider multiple branch lengths and different flow scenarios in addition to this baseline case.

We use CSMA as the MAC protocol from MUs to TAPs as well as from TAP to TAP and consider that MU-TAP and TAP-TAP transmission occurs on orthogonal channels. We simulate scenarios with 3, 5 and 20 MUs per TAP and depict the results for 5 MUs per TAP, as this factor has negligible impact on the aggregate throughput that we consider. Each MU is continuously backlogged and generates UDP traffic with 1000 byte packets. The channel rate is constant and has 2 Mb/sec capacity and there is no additional round trip time from the wired Internet. Moreover, the baseline scenario considers that TAPs two or more hops away are not in carrier sense range.

3.2 Results

Figure 3 depicts the received traffic at TAP_4, i.e., the goodput. The bars labelled "UDP/CSMA" represent the aggregate goodput for traffic originating at TAP_1, TAP_2, and TAP_3, along with the system-wide total (aggregate) goodput. Moreover, the bars labelled "Obj." depict the shares specified by the reference model in Section 2. Note that these shares represent an upper bound on achievable goodput subject to fairness constraints. While this objective considers an idealized scenario (perfect MAC, perfect scheduling etc.), it nevertheless serves as a benchmark for performance evaluation. The goodput of flows originating at TAP_4 are not shown here because this traffic does not interfere with TAP-TAP communication and therefore has no impact on system performance.

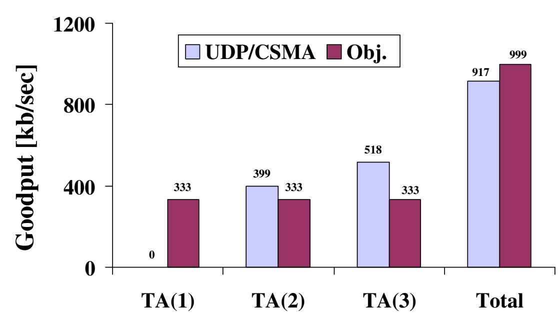


Figure 3: Goodput for Baseline Scenario

We make the following observations about these experiments. First, note that the obtained shares diverge significantly from the targets of the reference model, with traffic originating at TAP_1 *starved*. The starvation occurs because TAP_1 experiences exponential backoff far more frequently than TAP_2 and TAP_3 due to the "hidden terminal" problem [2]: when TAP_3 is sending to TAP_4, TAP_1 does not detect the transmission and attempts to acquire the channel. This results in collision at TAP_2 and exponential backoff for TAP_1. In Section 4.4.2, we consider a larger carrier sense range which results in TAP_1 traffic not starving.

Second, observe that flows originating an increased number of hops away from the wired destination contend for the channel an increasing number of times. This leads to a higher probability of collision and loss, and a corresponding throughput decrease.

Next, the right-most bars indicate that the system has achieved 92% of the *aggregate* capacity of the reference model. This high-

[2] TA(i) denotes the aggregate traffic flow ingressing at TAP i.

lights the importance of jointly considering capacity and fairness in performance analysis, as high aggregate capacity in the presence of starved flows is clearly undesirable.

Finally, note that because channel qualities have constant SNR in this scenario, there is no multi-rate transmission such that temporal fairness and throughput fairness are equivalent.

4. TCP FAIRNESS

In this section, we study the impact of the media access protocol, multi-rate channels, sector antennas, as well as different topologies and flow scenarios on TCP's fairness and capacity characteristics. In all experiments below, each MU generates long-lived TCP-Sack traffic, with all parameters set to their default values.

4.1 MAC, Hidden Terminals, and Asymmetry

Here, we modify the simulation scenario from the baseline by considering TCP-Sack and CSMA with and without Collision Avoidance. Figure 4 depicts the resulting goodput of the TAP-aggregated flows along with the targeted fair shares. As TCP acknowledgment packets form a traffic aggregate in the reverse direction, we depict this traffic as TAP_4's goodput.

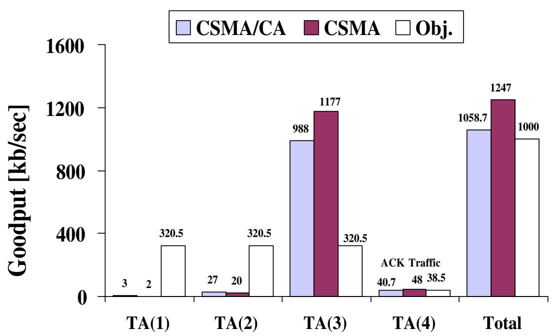


Figure 4: MAC and TCP Performance

Observe that neither the use of TCP nor the collision avoidance (CA) mechanism is able to prevent traffic originating at TAP_1 from starving. Moreover, TAP_2 traffic is now starved as well due to effects of TCP. The key reasons for this poor performance are (i) both TAP_1 and TAP_2 are now hidden terminals since TCP *acks* generate traffic in the opposite direction, [3] (ii) the mechanisms for increasing TCP window size generate bursts of packets that are mutually competing for the same medium, (iii) each TA flow consists of 5 sub-flows that are contending for the mobile to TAP channel, and (iv) losses can lead to timeouts, resulting in a congestion window of 1 segment and a significant throughput penalty.

These combined factors lead to starvation of traffic originating at TAPs 1 and 2. In the case of CSMA/CA, RTS/CTS exchange results in a decreased number of collisions, however introduces the problem of "information asymmetry" (see [2, 16, 24]), where TAP_1 has no information of transmissions between TAP_3 and TAP_4. This lack of information leads to TAP_1 backing off when not receiving a CTS from TAP_2 due to an ongoing transmission between TAP_3 and TAP_4. When the channel goes idle, TAP_3 or TAP_4 can immediately contend, whereas TAP_1 is in a backoff state. This problem is addressed via protocols such as RRTS [2] and DWOP [16]. However, in [32] the authors showed that RRTS cannot completely eliminate the problem.

Next, observe that CSMA obtains slightly higher goodput as compared to CSMA/CA as it does not incur overhead due to RTS/CTS exchange.

Finally, note that total goodput for TCP traffic is *higher* than that of the reference model objective. Starving multihop flows and giving all capacity to one hop flows is indeed the capacity-maximizing allocation.

4.2 Flow Scenario

Here, we study TCP's ability to exploit spatial reuse via the "Parallel Parking Lot" scenario depicted in Figure 5. The aggregate traffic from TAP_1, TA(1), consists of two sub-flows, flow (1,2) and flow (1,5), that have different egress TAPs, TAP_5 and TAP_2, respectively. Observe that transmission between the pairs TAP_1-TAP_2 and TAP_4-TAP_5 can occur simultaneously, allowing us to study TCP's spatial reuse capabilities. We perform experiments with both CSMA and CSMA/CA.

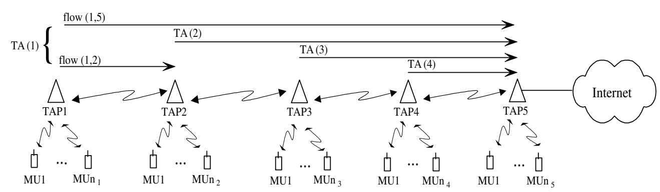


Figure 5: Parallel Parking Lot Scenario

Figure 6 depicts the results. For this scenario, the targeted fair shares are approximately 2/9 Mb/sec for all TAP-aggregated flows except for flow (1,2) that has fair share 6/9 Mb/sec (disregarding the *acks*). Observe that flows (1,5), TA(2), and TA(3) are starved for the reasons described above. While flow (1,2) is indeed able to exploit spatial reuse, it has done so only because TAP_3 traffic is starved: if this traffic was not starved, TAP_3 would be a hidden terminal for TAP_1 and would result in significant performance degradation even for the one-hop flow (1,2).

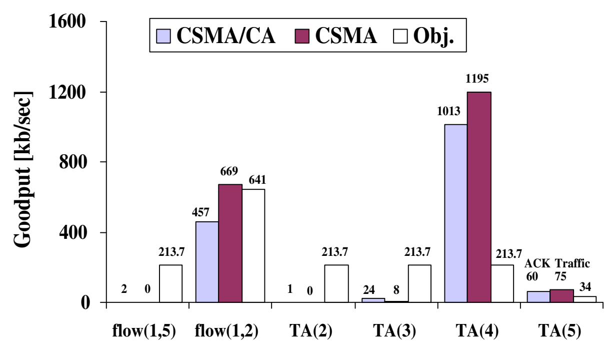


Figure 6: Parallel Parking Lot Results

4.3 Sector Antennas

Sector antennas provide statically-configured directional transmission and reception that results in increased spatial reuse and increased transmission range. Here, we consider that TAPs have

[3]Note that for the same reason TAP_4 is also a hidden terminal, however the number of *acks* from TAP_2 to TAP_1 is small such that the impact of this hidden terminal on the results is negligible.

sector antennas that spatially isolate downstream TAP transmissions from upstream TAP transmissions (i.e., inter-TAP links can be viewed as wireless point-to-point links). As in commercial systems, we also consider that each sector has its own air interface and MAC such that all sectors can be active simultaneously.

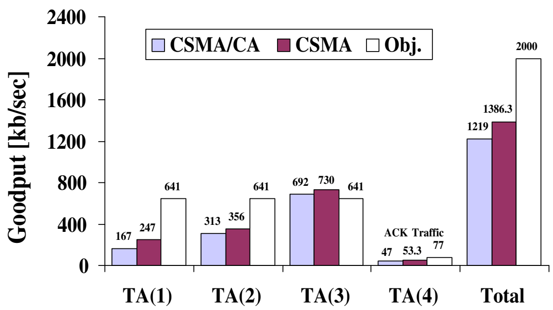


Figure 7: Sector Antennas

Figure 7 depicts the results. The reference model obtains throughputs of 641 kb/sec for each TAP: double that of the prior case due to the second antenna and MAC. Observe that the system's fairness properties are considerably improved with no flows being starved. Indeed, sector antennas have eliminated the hidden terminal problem and asymmetry problem discussed above.

However, the system still deviates significantly from the reference model for the following reasons. First, both MAC protocols result in a significant throughput bias for flows located fewer hops from the wired Internet. For example, TAP_1 traffic obtains 34% of the throughput of TAP_3, with the discrepancy increasing with path length. This occurs due to TCP's round-trip-time (RTT) bias, i.e., that TCP throughput is inversely proportional to RTT [25]. While propagation delay is negligible in these simulations, multi-hop flows incur increased contention and queueing as compared to one-hop flows.

Total goodput is increased as compared to the omnidirectional case because (i) TAP_1 and TAP_2 are no longer starved and now contribute to goodput, (ii) there is a second air interface, and (iii) there are reduced collisions. However, notice that the total goodput achieved by the system is 67% of that of the reference model for CSMA (and slightly less for CSMA/CA) and significantly less than twice that achieved with a single omnidirectional antenna. The key reasons for this limit are that TCP cannot perfectly utilize all available bandwidth and collisions are not *eliminated* as there is reverse traffic due to acknowledgment packets in addition to forward data traffic.

4.4 Channel Model

Finally, we consider variable rate channels and different propagation models.

4.4.1 *Multi-rate Transmission*

Here, we study variable channel conditions and MAC protocols that adapt their transmission rate according to SNR. In particular, we consider Receiver Based Auto Rate (RBAR) [13] and Opportunistic Auto Rate (OAR) [29]. Both use measured SNR of an RTS packet to set the transmission rate for the upcoming data packet in the CTS packet. RBAR targets throughput fairness by allowing one *packet* transmission per channel access, whereas OAR targets time-share fairness by allowing a maximum *time* duration per channel access.

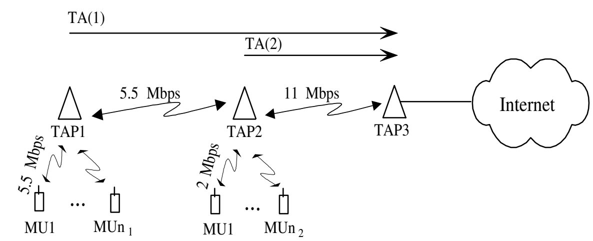


Figure 8: Multirate Channels

Figure 8 depicts the multirate scenario that we use in this set of experiments. The average channels qualities are 2, 5.5 or 11 Mb/sec as depicted in the figure.

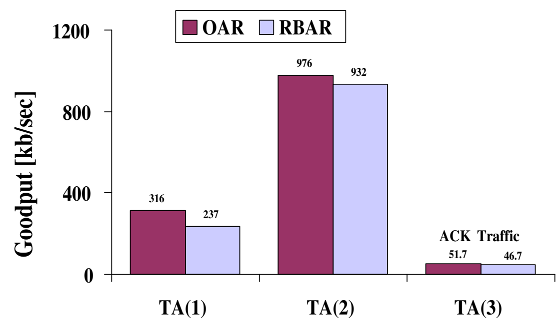


Figure 9: Multirate Transmission

Figure 9 depicts the goodput for each TAP aggregated flow measured as received traffic at TAP 3. Observe that even though the fairness characteristics are somewhat improved as compared to single-rate experiments, this occurs in part because channel qualities for TA(1) flows are considerably better than for TA(2) flows. Observe that the total goodput is still less than 1.7 Mb/sec for OAR, and less than 1.5 Mb/sec for RBAR, as TA(1) flows still suffer severe losses. However, in this scenario losses are limited to some extent due to TA(2) flows having poor channel quality (2 Mb/sec). Consequently, the traffic originating from TAP_2 is limited. On the other hand, the traffic from TA(1) is considerably increased. Lastly, we note that OAR results in slightly higher throughput than RBAR due to its use of consecutive packet transmissions under high-quality channel conditions.

4.4.2 *Impact of Carrier Sense Range*

While the above experiments consider scenarios in which the carrier sense range is less than twice the transmission range, here we study the impact of having a more sensitive carrier sense that results in a larger carrier sense range of twice the transmission range.

Figure 10 depicts the results for the Parking Lot scenario. Observe that the fairness characteristics are considerably improved, as the impact of hidden terminals and information asymmetry is mitigated, i.e., TAP_1 is aware of data transmission between TAP_3 and TAP_4. However, the total goodput is substantially reduced as compared to the case with smaller carrier sense range, since TAP_1 and TAP_2 are not starved and spatial reuse is inhibited.

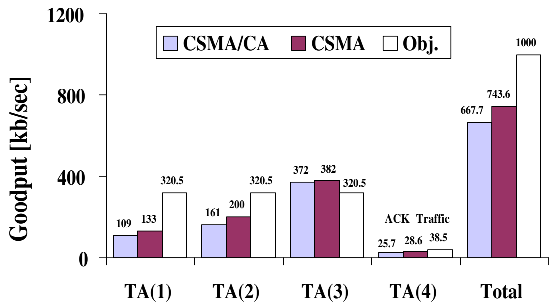


Figure 10: Impact of Carrier Sense Range

As the role of the carrier sense range was studied previously [30, 33], here we note that such a sensitive carrier sense range is not always realistic due to hardware limitations (see [33] for example). Nor is it necessarily desirable as a large carrier sense range also results in substantial reduction of spatial reuse and hence overall throughput. In any case, we note that an actual physical channel between any two nodes is under the influence of the scattering characteristics of the surrounding environment. Hence, the path loss between the nodes can be different in such a way that it results in a carrier sense range smaller than a preset value such as twice the transmission range.

5. INTER-TAP FAIRNESS ALGORITHM

The experiments presented in Sections 3 and 4 indicate that existing protocols incur severe unfairness and even flow starvation. In this section, we devise a distributed layer-2 protocol designed to achieve the objectives of the fairness reference model. Namely, the protocol attempts to eliminate the above starvation and unfairness by limiting flows at the first hop to their system-wide fair rate. The motivation for a layer 2 solution is that it does not require a special-purpose TCP for multi-hop wireless, it applies to UDP traffic, and it can react at faster time scales than end-to-end protocols. Moreover, in contrast to traditional congestion control techniques [20], our approach exploits the unique properties of multihop wireless backhaul networks such as TAP stationarity, the ability to treat branches independently, and limited path length to wires.

We next present the IFA algorithm, and then utilize it to simulate an idealized version of IFA in order to establish its baseline performance in simple scenarios and to study its interactions with TCP.

5.1 Inter-TAP Fairness Algorithm (IFA)

The objective of an Inter-TAP Fairness Algorithm is to allocate resources according to the TAP Reference Model via a distributed layer-2 protocol. The design space for IFA is immense as it encompasses not only classical congestion control issues encountered in wireline networks, but also issues unique to wireless networks (shared media, hidden terminals, fading channels, etc.) and unique to multihop backhaul networks (aggregate fairness granularity, removal of spatial bias, etc.).

Our focus here is not algorithm design itself as such an endeavor is beyond the scope of this paper. Rather, we describe an example specification of a layer 2 multi-hop wireless fairness algorithm that seeks to achieve the goals of the reference model. The algorithm has four key components described in the following protocol sketch.

Measurement of Offered Load and Capacity. Each TAP measures the average offered load for its own ingressing traffic, i.e., traffic arriving from its own mobile users or from its own wireline connection to the Internet. We denote the measured offered load of flow (i,j) as $\lambda^{(i,j)}$. This measurement can be performed at TAPs or MUs. If it is performed at TAPs, the measurement is noisy as TAPs will measure carried load vs. offered load. The measurement will be more accurate if performed by MUs themselves; however, such a realization would require messaging between MUs and TAPs such that TAPs can compute the aggregate offered load. We do not pursue this issue further as it is treated elsewhere, e.g., [24]. Each TAP also measures the average capacity of the links to each adjacent TAP. This capacity must include the effects of the MAC protocol, interference and multi-rate channels, hidden terminals, etc., and can be obtained via a combination of measurements and models via techniques such as [4, 34].

Message Distribution. Periodically, offered load and link capacities must be communicated to other TAPs. For each TAP$_i$ this message contains $\lambda^{(i,j)}$ for all $j \neq i$, and the capacities of all links to adjacent TAPs. Likewise, when aggregate fair shares are computed as a function of offered load, these shares must also be communicated. Note that as with measurement, this message distribution interval should be sufficiently small to track traffic and channel dynamics yet large enough to avoid excessive overhead. Thus, every T_{avg} seconds, each TAP sends a message to other TAPs containing its own offered load along with the computed fair shares and capacities of its links to the adjacent TAPs. Note that the signaling needs to be done among the TAPs within a branch only. Hence, as the number of TAPs per branch is small and the information can be represented efficiently (e.g., via a single byte per measurement or calculation), the algorithm's overhead is modest for averaging intervals beyond several hundred milliseconds. To illustrate with an example, denote the number of TAPs in a branch as N, and the number of links per TAP (i.e., degree) as d. Thus, for both forward and reverse message signaling, the total amount of overhead traffic is $N(N-1)(2d+N-1)/T_{avg}$ bytes/sec. For $N = 10$, $d = 2$, and $T_{avg} = 0.3$ sec, the overhead is less than 35 kb/sec. In any case, control messages require priority over data traffic and spare capacity should be reserved for control messages in order to ensure efficient operation of IFA.

Aggregate Fair Share Computation. A particular TAP has multiple links for which it is a sender or receiver, and each link is treated separately. Considering a particular link, the TAP computes the aggregate time shares in each of this link's contention neighborhoods and chooses the minimum value. The aggregate time shares are computed analogous to the fair share computation described in Section 6.2.2 and by using the offered loads from other ingress TAPs together with topology information and the link capacities of the contention neighborhood. Each TAP then converts the minimum time share to rate via use of the available link capacity and transmits it to other TAPs as described above. For example, consider the "parking lot" scenario shown in Figure 2 which will result in starvation of upstream TAPs under current protocols. As the reference model targets 1/6 of the capacity for each TAP-aggregated flow, the bottleneck link (between TAP$_3$ and TAP$_4$) can correctly compute the aggregate time share as 1/6 for this case of all flows being fully backlogged. A message containing the available rate for each ingress TAP (1/6 times the minimum link capacity on the flow's path) is then communicated upstream.

Ingress Rate Limiting. An ingress TAP will receive a TAP-aggregate fair-rate for each link as described above. Using these rates, the ingress TAP must determine its end-to-end fair flow rates, i.e., for ingress-egress flows. These rates are computed by treating the received TAP-aggregate rates as link capacities and by performing a computation similar to that presented in Section 6.2.3, with the difference that the extra "time capacity" is allocated to the flows able to exploit spatial reuse. There are two ways to implement rate limiting. First, the TAP can signal each MU of its fair share and the MU can use a rate limiter such as a leaky bucket to realize the share. Second, the TAP can CTS or poll MUs at the desired rate in order to achieve the targeted share.

Thus, by enforcing MUs to throttle at their ingress point to this system-wide fair rate, the algorithm targets to achieve the reference model's objectives. Moreover, by achieving the fair rates at layer 2, TCP needs only to "fill the pipe," a far simpler task than attempting to track the fair rates itself. Thus, IFA targets to achieve the reference model for UDP or TCP-controlled traffic.

5.2 IFA Simulation

We simulate an idealized version of the IFA algorithm in which the local-fair share computation is based on perfect information of the offered load and the link capacities. In practice, this information would be obtained as described above. Next, if not otherwise mentioned, we set the averaging interval to $T_{avg} = 0.3$ sec and send a message to neighboring TAPs of the average "true" offered load every averaging interval. Finally, we implement the rate limiters at MUs.

5.3 Baseline

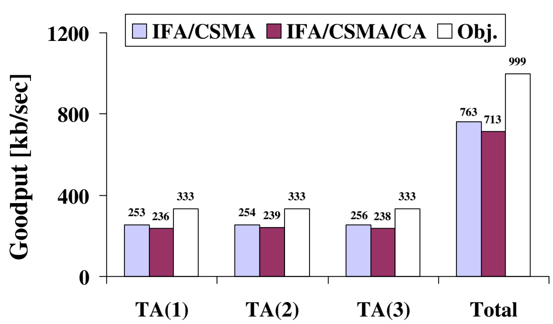


Figure 11: UDP/IFA with CSMA and CSMA/CA

Here, we study the performance of continuously backlogged UDP flows and CSMA as well as CSMA/CA media access as implemented in IEEE 802.11. The results are depicted in Figure 11 which shows the aggregate traffic from each TAP under IFA and the reference model. Observe that IFA/CSMA achieves a nearly identical goodput for each TAP of 253 kb/sec to 256 kb/sec, despite the presence of hidden terminals and use of the CSMA MAC. However, IFA does not eliminate the hidden terminal problem, and for other traffic matrices CSMA/CA is required. In any case, by controlling the input rate of TAP_3, TAP_1 and TAP_2 are able to access the channel and achieve their fair shares. Time limiting TAP_3 to transmit only $1/6^{th}$ of the time, significantly reduces link layer contention, and provides sufficient spare capacity for the hidden terminal TAP_1.

Likewise, CSMA/CA also attains near equal throughput for each TAP at 236 to 238 kb/sec, approximately 7% less than that achieved by CSMA primarily due to RTS/CTS overhead.[4]

On the other hand, IFA over CSMA and CSMA/CA respectively achieves 76% and 71% of the per-TAP and aggregate throughput as compared to the idealized reference model. This discrepancy occurs due to imperfect media access and collisions. Namely, while in this case all flows are indeed throttled to their exact ideal system-wide fair time shares, collisions and retransmissions occur due to MAC layer contention resulting in less than 100% efficiency. In any case, as 76% to 71% goodput is in the range of the maximum achievable by IEEE 802.11 as indicated by models and simulation studies [3], achieving higher performance would require reduced collisions or other MAC enhancements.

Thus, the results indicate that by throttling *input* traffic to its system wide fair time share, even severe MAC problems such as high loss due to hidden terminals and contention can be alleviated and the network's fairness objectives can be approximately achieved.

5.4 Spatial Reuse

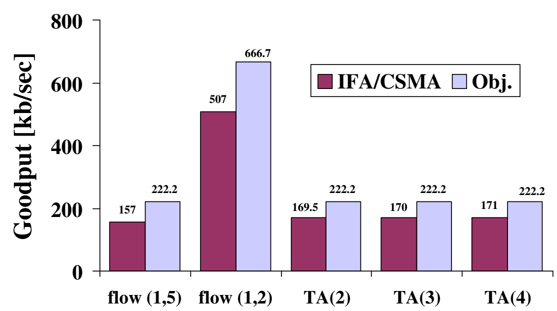


Figure 12: IFA in the Parallel Parking Lot

Next, we study IFA's ability to exploit spatial reuse. We consider the scenario depicted in Figure 5 with continuously backlogged UDP flows and present the results in Figure 12. We observe that overall performance is considerably improved as compared to that obtained by using TCP without IFA. Moreover, IFA is able to exploit spatial reuse for the one hop flow (1,2), since transmission between TAP_1 and TAP_2 is in different contention neighborhood as compared to transmission between TAP_4 and TAP_5. However, observe that the throughput share of flow (1,5) is almost 10% less than the shares of TA(2), TA(3), and TA(4). The reason is that TAP_1 has a larger number of MUs (flows (1,2) and (1,5)) in its collision domain as compared to TAP_2, TAP_3, and TAP_4. Thus, the increased contention reduces this TAP's flows' ability to utilize all of their available resources.

5.5 TCP/IFA

While the above experiments demonstrate the potential of a layer 2 fairness algorithm, end-to-end congestion control at layer 4 is still required in case the bottleneck is not in the multi-hop wireless backhaul network. Thus, here we consider interactions of IFA and TCP. The scenario consists of MUs generating long-lived TCP Sack flows in the Parking Lot scenario from Figure 2.

The aggregate TCP goodput for flows originating at different TAPs is depicted in Figure 13 along with the target bandwidth shares. As shown, TCP's end-to-end performance is considerably

[4] Note that in most commercial implementations, RTS/CTS is disabled by default, as it is enabled only for packets above a threshold that is set to the maximum 1500 bytes by default.

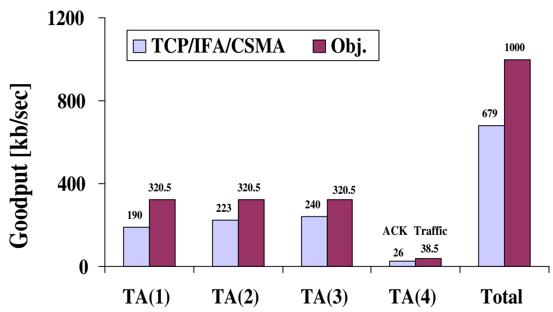


Figure 13: TCP over Inter-TAP Fairness Algorithm

improved by the IFA algorithm (cf. Figure 4) with TAPs 1 to 3 obtaining throughput that is 59% to 75% of the objective function and an aggregate capacity of 65% of the objective function (a reduction of 14% as compared to UPD/IFA). IFA has improved the performance of TCP for the same reasons as with UDP traffic (reduction of contention losses, etc.) which prevents TCP from incurring excessive window decreases and timeouts, and prevents starvation of traffic from TAPs 1 and 2. Moreover, with TCP over IFA, TCP cannot inject bursts of packets in the network, so that the occurrence of excessive losses and timeouts are eliminated as TCP traffic is smoothed by the use of rate controllers.

On the other hand TCP does introduce an increased spatial bias as it favors short RTT flows: IFA alone cannot completely counter this effect. Likewise, reclaiming the 14% throughput loss of TCP/IFA as compared to UDP/IFA would likely require an enhancement to TCP via techniques such as those described in Section 7.

5.6 Inter-TAP Performance Isolation

The unfairness between congestion responsive TCP and nonresponsive constantly backlogged UDP flows is well established. Yet, the objective of the IFA protocol is to provide inter-node performance isolation, *independent* of the traffic types. That is, if traffic originating at one TAP is continuously backlogged UDP traffic, TCP traffic originating at upstream or downstream TAPs should not be penalized.

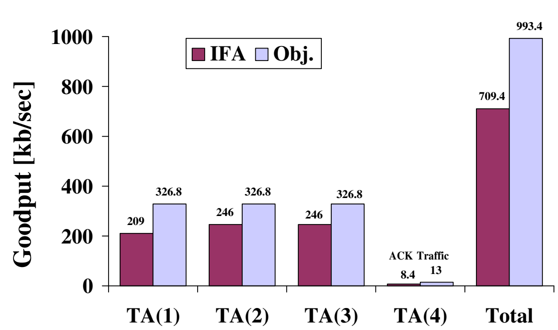


Figure 14: Performance Isolation for TCP Traffic

To explore this scenario, we consider the Parking Lot in which each TAP has one MU, and the MU from TAP_1 transmits TCP traffic, while MUs from TAP_2 and TAP_3 transmit continuously-backlogged UDP traffic. The results are depicted in Figure 14. Observe that the TCP flow obtains 64% of the idealized objective throughput, whereas the UDP flows obtain 75%. Thus, by throttling uncontrolled UDP flows at the input, IFA ensures that an upstream TCP flow can obtain nearly its fair share, with the differences between TCP and UDP shares in isolation explored previously. However, we do note that having an increased number of flows in each TAP's collision domain would result in a slight degradation of goodput for TCP flows, as in the presence of even balanced contention loss, TCP flows reduce their rate, whereas UDP rates remain the same.

5.7 Unbalanced Flows

The TAP reference model defines fairness at the TAP-aggregated granularity, meaning that each TAP-aggregated flow should achieve the same time share regardless of the number of mobile users in its collision domain. In this final set of experiments, we modify the number of mobile users per TAP, such that TAP_1 and TAP_2 each have two MUs transmitting constant-rate UDP traffic, whereas TAP_3 has only one MU transmitting TCP traffic. We study the ability of the IFA protocol to provide TAP-aggregated fairness as opposed to per flow fairness.

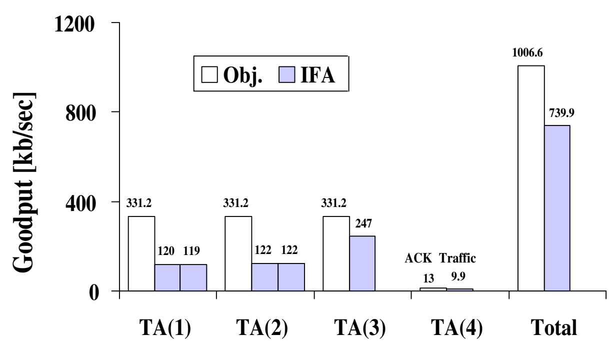


Figure 15: Unbalance Number of Flows per TAP

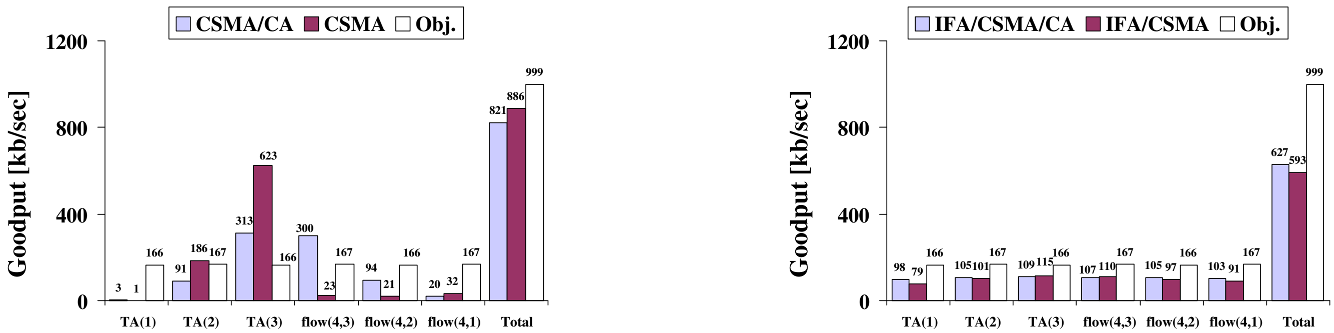
The results are depicted in Figure 15 which illustrates that each mobile user from TAP_1 and TAP_2 obtain nearly the same goodput which is approximately half of that obtained by the TCP flow from TAP_3. This is indeed the targeted behavior as each TAP attempts to provide a service analogous to that achieved by an access point with a *wired* backhaul link.

5.8 Forward and Reverse Traffic

Here, we consider a complete traffic to and from the wired TAP as depicted in Figure 16. In this scenario, it is desirable to employ *weighted* fairness in order to ensure that forward and reverse traffic have the same throughput. Hence, because TAP_4 traffic consists of all reverse (or downlink) traffic, it requires a higher weight than TAPs 1 to 3.

We simulate a scenario as in the figure with all link capacities set to 2 Mb/sec. The weight of TAP_4 is $w_4 = 3w_i$, $i = 1, 2, 3$ so that the fair shares under weighted fairness are TA(i)=flow(4,i)=2/12 Mb/sec, $i = 1, 2, 3$. Note that with equal weights, the shares would be TA(1)=TA(2)=TA(3)=2/8 Mb/sec, whereas each flow in TAP aggregate TA(4) (flow (4,3), flow (4,2), and flow (4,1)) would have a fair share of 2/24 Mb/sec.

Figure 17 depicts the results for UDP and UDP/IFA with CSMA



(a) UDP with CSMA and CSMA/CA

(b) UDP/IFA with CSMA and CSMA/CA

Figure 17: Goodput for Forward and Reverse Traffic

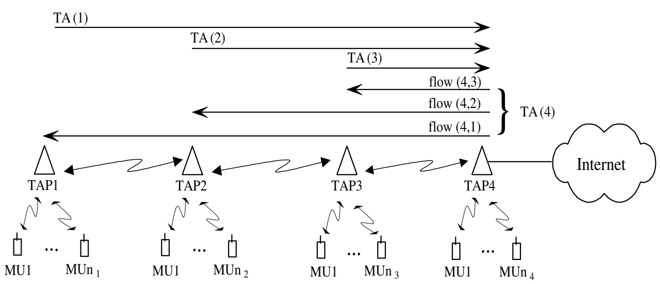


Figure 16: Scenario with Forward and Reverse Traffic

and CSMA/CA. Observe that for CSMA, downlink traffic is considerably lower as compared to uplink traffic and contributes to total goodput with only 76 kb/sec, whereas uplink traffic contributes with 810 kb/sec. This occurs due to compounded effects of hidden terminals. Because RTS/CTS mitigates effects of hidden terminals, the results for CSMA/CA are more balanced, and downlink traffic contributes with 414 kb/sec whereas uplink traffic contributes with 407 kb/sec. Yet, as found previously, the longest path traffic has considerably lower goodput as compared to the shorter path traffic.

In contrast, IFA achieves significantly better fairness properties and, as targeted, TA(4) is close to three times TA(i), $i = 1, 2, 3$. On the other hand, the total throughput is lower (9% for CSMA/CA and 13% for CSMA) for this scenario as compared to the scenario with forward traffic only. The reason for this is the decreased available bandwidth caused by increased contention due to the existence of reverse traffic and an increase in the number of transmitting nodes.

6. CAPACITY AND FAIRNESS

In this section we evaluate the effect of fairness constraints on the capacity of a multi-hop wireless backhaul network. Specifically, we first compute the maximum aggregate throughput when there is no fairness constraint and show that it results in starvation of some flows in a multihop scenario with multiple flows. Then we study the effect of each of the fairness objectives on the achievable throughput of aggregate flows in the network.

While capacity of multi-hop wireless networks has received significant attention, e.g., [9, 10, 11, 15, 21], the critical relationship between fairness and throughput in multi-hop networks has been largely unstudied. Instead the focus has been on analyzing and maximizing the achievable aggregate throughput in multi-hop networks. In [15], the authors formulate the throughput maximization problem in a general setting and find upper and lower bounds of throughput for any network topology and system parameters. However, the formulation in [15] lacks fairness constraints and seeks to maximize the number of packets leaving the source and arriving at the destination. Consequently, in a multi-flow scenario, this can lead to significant unfairness and result in starvation of some flows. By applying the fairness constraints to the LP formulation developed in [15], one can use the methodology of [15] to find fairness-constrained bounds on throughput. However, the constraints of the fairness reference model, i.e., the fairness objectives, affect the solution space of the optimization problem. In the case of the TAP fairness reference model, the ingress-aggregation constraint results in non-linear constraints, and consequently significantly increases the complexity of the solution. In this work, however, we consider scenarios applicable to wireless backhaul networks that result in a reduced and computationally feasible solution space.

The problem of finding network throughput, i.e., the solution to the LP problem, is topology dependent. It depends not only on the nodes, but also on the number of flows and their routes. Here we consider the specific characteristics of a multihop wireless backhaul network and target to derive a general formulation to compute and evaluate its throughput, with and without fairness constraints. Hence, to be able to separate the effect of fairness constraints from the effect of topology and spatial reuse on the throughput of the network, we focus on the throughput in a network in which no spatial reuse is possible, i.e., only one link can be active at any given time. This is specifically true in a clique, a region in which all links mutually contend. We consider a multihop network with a perfect collision free MAC and a fluid arrival and service model. The network consists of N nodes and F flows. Each flow f traverses one pre-determined route r_f, with the number of hops (or wireless links) flow f traverses denoted by h_f. Each wireless link, l, has a fixed capacity, C_l, which is a function of the distance between the transmitter and the receiver pair and the surrounding environment. Moreover the mobile users use a different wireless channel from the inter-TAP links and hence their contention neighborhoods are different. We consider the aggregate inter-TAP flows to be always backlogged and focus on the TAP-aggregate throughput.

6.1 Aggregate Throughput without Fairness Constraints

With a system free of fairness constraints the goal is to assign flow time shares and rates such that network throughput is maximized.

Let ρ^f denote the long-term throughput of flow f. Then the time shares that maximize the aggregate network throughput provide a solution to the following optimization problem,

$$\begin{aligned} \Gamma &= \max_{\{t_l^f\}} \sum_{f=1}^{F} \rho^f \\ & \text{s.t.} \sum_{f=1}^{F} \sum_{l \in r_f} t_l^f \leq 1, \end{aligned} \tag{1}$$

where $\rho^f = \min_{l \in r_f} t_l^f C_l$ and t_l^f denotes the time share of flow f at link l, $l \in r_f$. In Equation (1), the maximum aggregate throughput is achieved only if

$$\sum_{f=1}^{F} \sum_{l \in r_f} t_l^f = 1. \tag{2}$$

Otherwise there exists spare "time capacity" which results in reduction of throughput, as we are considering a clique. Moreover, to achieve maximum throughput, the time share assigned to each link must be such that flow preservation properties are satisfied, i.e., the time share must be equal to the time required for forwarding all incoming packets. If it is shorter, there are packets that have been transmitted by previous links but cannot get to the destination, which indicates that the time used by the previous links to transmit those packets was not efficiently used. On the other hand, if this time is longer, the link will be idle during part of the allocated time share. Hence,

$$t_i^f C_i = t_j^f C_j, \quad \forall i, j \in r_f. \tag{3}$$

A solution to the maximization problem of (1) is to assign time shares such that only the flow with the maximum throughput is allocated time to transmit and all other flows are starved. Then, the maximum aggregate throughput, Γ, is given by

$$\Gamma = \max_{1 \leq f \leq F} \bar{\rho}^f, \tag{4}$$

where $\bar{\rho}^f$ is the throughput of flow f when f is the only flow which has been allocated time to transmit, i.e.,

$$\sum_{l \in r_f} t_l^f = 1. \tag{5}$$

By solving Equations (3,5), $\bar{\rho}^f$ is computed as

$$\bar{\rho}^f = \left(\sum_{l \in r_f} \frac{1}{C_l} \right)^{-1}. \tag{6}$$

Thus, the time shares of flow f to achieve $\bar{\rho}^f$ are $t_l^f = \frac{\bar{\rho}^f}{C_l}$, for $l \in r_f$.

If there is more than one flow in the network which satisfies the maximum throughput requirement, the time-share/rate assignment problem will have multiple solutions in which all provide maximum throughput. It is, however, important to note that in all cases, throughput is maximized in exchange for possible starvation of multiple flows in the network.

6.2 Aggregate Throughput with Fairness Constraints

The achievable throughput in a wireless network is dependent on the strictness of the fairness constraints in the system, as each additional fairness constraint can potentially reduce the throughput of the system. As described in the appendix, TAP fairness requires satisfaction of four main objectives, *temporal fairness*, *spatial bias*, *ingress aggregate*, and *spatial reuse*. Each of these objectives, as described in the appendix, imposes additional system constraints. In the rest of this subsection, we study the individual effect of these constraints and use several examples to illustrate their effect on aggregate throughput. As in the previous subsection, we consider a clique, where spatial reuse is not possible.

6.2.1 Temporal Fairness Constraint

With temporal fairness, channel access *time*, rather than bandwidth, is considered to be the system resource. Under temporal fairness constraints, the total time that any flow is active is equal for all flows, regardless of the number of hops they traverse and their link capacities. Hence, the time shares are a function of the number of flows in the network. The effect of temporal fairness on throughput for single-hop flows has been studied in [29]. Here, we study the effect of temporal fairness on throughput of a multi-hop backhaul network.

Consider flow f traversing route r_f. Then, the total time share of a flow f is the sum of time shares of its individual links, which, based on the temporal fairness constraint, is equal for all flows. Hence,

$$\sum_{l \in r_f} t_l^f = \frac{1}{F}. \tag{7}$$

Equation (7) adds an additional constraint to the optimization problem of (1). Equations (3, 7) represent a system of linear equations in which the number of equations and unknowns are identical and equal to $\Pi_{1 \leq f \leq F} h_f$. The solution of this system yields the time share of flow f over link l,

$$t_l^f = \frac{1}{F C_l \sum_{i \in r_f} \frac{1}{C_i}}. \tag{8}$$

The throughput of flow f is then

$$\begin{aligned} \rho^f &= \frac{1}{F \sum_{i=1}^{h_f} \frac{1}{C_i}} \\ &= \frac{\bar{\rho}^f}{F}. \end{aligned} \tag{9}$$

This result follows the intuition that under temporal fairness, each flow receives $1/F$ of the share of throughput it would have received if it was the only contending flow in the network.

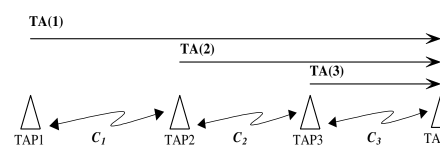


Figure 18: Example scenario for throughput computation under temporal fairness constraint

Consider the example shown in Figure 18. Flows TA(1), TA(2), and TA(3) represent the aggregate traffic from the respective TAPs. Consider C_1, C_2, and C_3 to be respectively 20 Mb/sec, 5 Mb/sec, and 10 Mb/sec. With temporal fairness, these flows receive equal transmission times. Defining the vector $T_l = [t_l^{TA(1)}, t_l^{TA(2)}, t_l^{TA(3)}]$, the time shares of the three flows of the three links are computed to be $T_1 = [\frac{1}{21}, 0, 0]$, $T_2 = [\frac{4}{21}, \frac{2}{9}, 0]$, and $T_3 = [\frac{2}{21}, \frac{1}{9}, \frac{1}{3}]$. Then the normalized throughput of flows TA(1), TA(2), and TA(3) is computed as 18%, 21%, and 61%, respectively. Note that as the number of hops for a flow increases its throughput decreases. That is because the common resource should be shared among the links of the flow. Next we add the spatial-bias-removal constraint which prevents throughput reduction of flows that traverse multiple hops.

6.2.2 *Spatial Bias Constraint*

Allocating equal shares of the system resource to different flows in a network may result in *spatial bias*, i.e., flows with a smaller number of hops receive higher throughput compared to flows that traverse a larger number of hops and must share allocated resource on all links on the path. Under the spatial-bias-removal constraint, the *fair share* of all flows, as described in the appendix, should be equal, i.e.,

$$T_a(f) = T_a(g) \quad \text{for all flows } f \text{ and } g, \tag{10}$$

where $T_a(f)$ is the *fair share* of flow f, defined in the appendix. Equation (10) results in $F-1$ equations, which in addition to Equations (2, 3) results in a linear system with a unique solution. Here we solve this system of equations for the system policy presented in the appendix. Restating the policy expressed in Equation (28) we have

$$T_a(f) = t^f_{l^f_1} \quad \text{for any flow } f. \tag{11}$$

Then, Equation (10) can be rewritten as

$$t^f_{l^f_1} = t^g_{l^g_1}, \quad \text{for all flows } f \text{ and } g, \tag{12}$$

where l^f_1 denotes the first link in the route of flow f. Solving Equations (2, 3,12), the time share of any flow of its first hop is computed as

$$t^i_{l^i_1} = \left(\sum_{f=1}^{F} \frac{C_{l^f_1}}{\bar{\rho}^f} \right)^{-1}, \quad 1 \leq i \leq F. \tag{13}$$

The throughput of flow f is then

$$\rho^f = t^f_{l^f_1} C_{l^f_1}, \tag{14}$$

and its time share for any link l, $l \in r_f$, is

$$t^f_l = \frac{\rho^f}{C_l}. \tag{15}$$

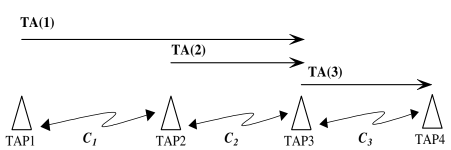


Figure 19: Example scenario for throughput computation under the spatial-bias-removal constraint

Consider the scenario depicted in Figure 19 with link capacities C_1, C_2, and C_3 set as in the example of Figure 18. With temporal fairness and spatial-bias-removal constraints, the throughput of each of the flows TA(1), TA(2), and TA(3) would be 2.85 Mb/sec, 1.43 Mb/sec, and .71 Mb/sec, respectively. Note that under the policy of Equation (11) the throughput of a flow is proportional to the capacity of its ingress link.

6.2.3 *Ingress Aggregation Constraint*

In computing flow throughput with the ingress aggregation constraint, multiple flows which initiate at the same ingress node are treated as one. In this subsection, we first compute the aggregate throughput for a system with only *temporal fairness* and *ingress aggregate* constraints, and then add the *spatial-bias-removal* constraint to the problem.

Consider n_a aggregate flows originating from a single TAP.[5] Let A denote the set of n_a aggregate flows. With temporal fairness and ingress aggregation constraints, the set of aggregate flows receives the same time share of any other flow in the network. Hence, the total time share of A is equal to $1/F$ and for each aggregate flow f_a,

$$\sum_{l \in r_{f_a}} t^{f_a}_l = \frac{1}{n_a(F - n_a + 1)}. \tag{16}$$

For any other flow f in the network,

$$\sum_{l \in r_f} t^f_l = \frac{1}{F - n_a + 1}. \tag{17}$$

Solving Equations (3, 16, 17) yields the throughput of flows in the network. For any aggregate flow f_a,

$$\rho^{f_a} = \frac{\bar{\rho}^{f_a}}{n_a(F - n_a + 1)}, \tag{18}$$

and for all other flows

$$\rho^f = \frac{\bar{\rho}^f}{(F - n_a + 1)}. \tag{19}$$

We next add the spatial-bias-removal constraint. Without loss of generality, we assume that aggregate flows are numbered 1 to n_a. Then, based on the spatial-bias-removal constraint, as presented in Section 6.2.2, for any two non aggregate flows, f and g,

$$T_a(f) = T_a(g) \quad n_a < f \leq F \text{ and } n_a < g \leq F. \tag{20}$$

Similarly, for any two aggregate flows, f_a and g_a,

$$T_a(f_a) = T_a(g_a) \quad 1 \leq f_a \leq n_a \text{ and } 1 \leq g_a \leq n_a. \tag{21}$$

Moreover, because of the ingress aggregate constraint

$$\begin{aligned} T_a(A) &= \sum_{f=1}^{n_a} T_a(f_a) \\ &= T_a(g) \quad n_a < g \leq F. \end{aligned} \tag{22}$$

Equations (20, 21, 22) provide $F-1$ equations. Similar to the previous section, these $F-1$ equations with Equations (2, 3) form a system of linear equations with a unique solution.

Here we solve the resultant system for $TA(\cdot)$ defined in Equation (28) and restated in Equation (11). Substituting $TA(\cdot)$ as expressed in Equation (11) in Equations (20, 21, 22), and solving the linear system formed by these equations and Equations (2, 3), the throughput of any non aggregate flow g and aggregate flow g_a is computed as

$$\rho^g = \frac{C_{l^g_1}}{\frac{1}{n_a}\sum_{f_a=1}^{n_a} \frac{C_{l^{f_a}_1}}{\bar{\rho}^{f_a}} + \sum_{f=n_a+1}^{F} \frac{C_{l^f_1}}{\bar{\rho}^f}}, \tag{23}$$

and

$$\rho^{g_a} = \frac{C_{l^{g_a}_1}}{\sum_{f_a=1}^{n_a} \frac{C_{l^{f_a}_1}}{\bar{\rho}^{f_a}} + n_a \sum_{f=n_a+1}^{F} \frac{C_{l^f_1}}{\bar{\rho}^f}}, \tag{24}$$

respectively. Having computed the throughput of any flow f, its time share of any link l, $l \in r_f$ is given by Equation (15).

[5]Generalization to multiple TAPs is straight forward.

Next we consider different subsets of fairness objectives and compare the aggregate throughput under the constraints imposed by these objectives.

6.3 Throughput Comparison under Different Fairness Objectives

Consider the scenario depicted in Figure 20. We keep the link capacities as set in the example of Figure 18. The throughput of each flow as well as the total throughput for different fairness constraints is shown in Table 1.

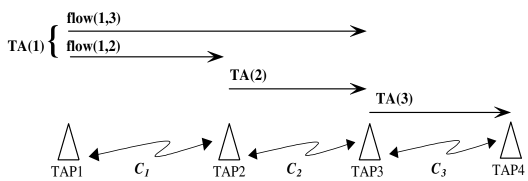


Figure 20: Example scenario for throughput comparison under different fairness constraints

Without any fairness constraints, the maximum throughput is achievable by assigning 100% share to flow (1,2) and starving all others. This flow traverses only one hop which has a capacity of 20 Mb/sec. Adding temporal fairness constraints to the system, all four flows would have equal access time to the channel. This equal time share policy results in higher throughput for flows flow (1,2) and TA(3), which also traverse only one hop. Note that although both flows have equal transmission time, the throughput of flow (1,2) is twice as high as the throughput of flow TA(3) due to its access to a higher quality channel allowing a higher transmission rate.

Under spatial-bias-removal and throughput constraints, we require equal throughput for all flows. This results in a higher time share for multi-hop flows and links with lower capacity, and consequently, in lower aggregate throughput. The spatial-bias-removal constraint with the temporal-fairness constraint provide all flows with equal time share on their ingress links.[6] Hence, because C_1 is twice as high as C_3, which in turn is twice as high as C_2, flow (1,3) and flow (1,2) are each capable of achieving a higher throughput compared to TA(3), which itself is achieving a throughput twice as high as TA(2)'s.

With the ingress aggregation requirement and either throughput or temporal fairness implemented in the system, aggregate flows (1,3) and (1,2) are required to share resources and hence achieve a smaller throughput. Moreover, since flow (1,3) is a multihop flow, its throughput is significantly lower as compared to flow (1,2).

Requiring flows originating from an ingress TAP to share resources assigned to the aggregate flow under the spatial-bias-removal constraint and either *throughput* or *temporal* fairness reduces the share for ingress aggregate flows and fairly assigns it to the other flows. Under both sets of constraints the ingress flows receive equal shares. Under *throughput* fairness, this results in equal throughput for aggregate flows TA(1) (flows (1,2) and (1,3)), TA(2), and TA(3). Under *temporal* fairness, however, a difference among throughput of different flows is highly probable depending on their link capacities.

Finally, observe that proportional fairness penalizes multi-hop flows. Moreover, TAP fairness provides a total throughput close to that of proportional fairness.

[6]The relationship between the time shares of different flows under the spatial-bias-removal constraint is determined by the system policy described in the appendix.

Table 1: Comparison of aggregate throughput of scenario depicted in Figure 20 for different fairness constraints.

Fairness Constraints	Throughput (Mbps)				
	flow(1,3)	flow(1,2)	TA(2)	TA(3)	Total
None	0.00	20	0.00	0.00	20
Temporal	1.00	5.00	1.25	2.50	9.75
Spatial Bias & Throughput	1.66	1.66	1.66	1.66	6.64
Spatial Bias & Temporal	2.50	2.50	.625	1.25	6.87
Ingress Agg. & Throughput	0.64	1.28	2.56	2.56	7.04
Ingress Agg. & Temporal	0.66	3.33	1.66	3.33	8.98
Ingress Agg. & Spatial Bias & Throughput	1.25	1.25	2.50	2.50	7.50
Ingress Agg. & Spatial Bias & Temporal	2.00	2.00	1.00	2.00	7.00
Proportional	1.05	2.10	2.10	2.10	7.35

Under the TAP fairness reference model of Section 2, the fairness objectives include temporal fairness, removal of spatial bias, and ingress aggregation; it additionally requires maximization of spatial reuse in the system.

In a clique, where no spatial reuse is possible, the throughput of a flow can be computed using the formulation provided above. In a general network, however, where simultaneous transmissions due to existence of spatial reuse is possible, the sum of all time shares in the network can exceed one. Taking this fact into account and having the routing information, a similar approach to the one presented in this section can be used to compute the shares of different flows.

Thus, as summarized in the table, any of the fairness constraints restricts system capacity to less than its maximum value for this example. Yet, the rate allocations for the throughput maximizing case result in starvation making it clearly undesirable. With fairness constraints, temporal fairness significantly increases capacity as compared to throughput fairness as demonstrated for single-hop networks in [29, 12] and generalized to multi-hop networks here. Likewise, removing spatial bias has its own capacity cost yet may be essential for multi-hop backhaul networks in order to provide access links independent of spatial location, as is the case in wireline access networks. Finally, as TAPs correspond to administrative domains (hot spots, residences, etc.), it is critical to have TAP vs. flow fairness as the latter would make one TAP's performance vulnerable to another TAP having many micro-flows.

7. RELATED WORK

Fairness in ad hoc networks has been the focus of intense research efforts, e.g., [12, 14, 24, 22, 28, 31, 35]. While [14, 22, 31] aim at implementing max-min fairness, [24, 35] address proportional fairness. Reference [28] shows that max-min fairness in an ad hoc wireless network without battery-life constraints results in equalizing all rates to the smallest rate flow. Moreover, IEEE 802.11 aims to achieve max-min fairness and hence results in the performance anomaly presented in [12]. Likewise, the maximizing capacity formulation in [28] illustrates our modeling result in Section 6: without fairness constraints, network efficiency is achieved at the expense of starvation of low-capacity flows. Thus, [28] proposes proportional fairness as a trade-off between efficiency and fairness in a wireless ad hoc network. In this work, however, we focus on the unique characteristics of a wireless backhaul network and define a fairness model that addresses the requirements of multihop aggregated flows.

Previous work targeting realization of different fairness reference models has mostly been focused either on single-hop flows, e.g., [12, 24] or on TCP modifications (discussed below). An exception is reference [35], which makes a case for the feasibility of hop-by-hop schemes in ad hoc networks and proposes a distributed layer 2 congestion control mechanism that aims at achieving per-flow proportional fairness [18]. Thus, like IFA, [35] addresses multihop fairness via a distributed layer 2 protocol. In contrast, IFA targets time share fairness, removal of spatial bias, etc. as targeted by the reference model. Moreover, our focus here is not on algorithm design itself, but rather on performance analysis of the solution space.

Performance of TCP over wireless networks with IEEE 802.11 media access has been studied extensively. In [19], a throughput analysis of a multihop chain topology with a simplified MAC is performed. In [27], TCP fairness in a single-hop wireless LAN is studied to show the effect of the access point's queue size on unfairness towards downlink flows. The effect of TCP congestion window limits on performance of multihop networks has been evaluated in [6]. The results analytically confirm previous simulation-based solutions indicating that a smaller congestion window limit will improve the performance of TCP over multihop wireless paths. Upper bounds on the bandwidth-delay-product in a multihop path are provided along with an adaptive strategy that dynamically adjusts the TCP congestion window limit to ensure that it will not exceed the bound. Reference [7] also focuses on congestion window limits for TCP and shows that for any topology there is an optimal window size that maximizes TCP throughput. This optimal size is computed for a chain to be the number of hops divided by four. In [32], TCP fairness among flows in an ad hoc network connected to the Internet through gateway nodes is studied. Simulations and testbed measurements are used to compare multihop flows to single-hop flows, and wireless-to-wired flows to wired-to-wireless flows, and to study the subsequent unfairness. The authors conclude that hidden terminals are the main reason for unfairness in TCP and reference [7] indicates that significantly higher throughput is achievable by increasing the carrier sense range. In contrast, we have found that rate-limiting downstream flows to their fair rate can alleviate the effects of hidden terminals upstream.

Unlike prior work on TCP over multihop wireless networks, we employ a fractional factorial experimental design to identify the joint performance factors that lead to poor performance, unfairness, and flow starvation. In addition, we focus on the unique objectives of multihop wireless backhaul networks. Lastly, while IFA could potentially be integrated into TCP, we note that only a layer 2 solution can protect TAPs from non-responsive UDP flows as demonstrated in Section 5.6, and only a layer 2 solution will not require significant modification to TCP.

Finally, we note that ingress-aggregated fairness as defined in [8] is identified as the objective for IEEE 802.17, a protocol for metropolitan networks.[7] However, as IEEE 802.17 is a wireline protocol, our framework here is quite different as we incorporate variable rate channels, shared media access, etc.

[7]http://ieee802.org/17

8. CONCLUSIONS

Multihop wireless backhaul networks have the potential to provide economically viable broadband access networks. Unfortunately, we have shown that current protocols can result in severe unfairness and even starvation of flows farther away from wired Internet entry points. We developed an idealized reference model that characterizes the unique performance objectives of multihop wireless backhaul networks. We performed extensive simulation experiments to identify the individual and joint performance factors that lead to performance problems. We developed a simple distributed layer two fairness algorithm to demonstrate the possibility of achieving fairness objectives without modification to TCP, and in the presence of non-TCP non-responsive flows. We studied the relationship between fairness and system capacity and quantified the cost of fairness, as compared to a capacity-maximizing strategy which can starve multihop flows.

9. REFERENCES

[1] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, 1992.

[2] V. Bharghavan, S. Demers, S. Shenker, and L. Zhang. MACAW: A media access protocol for wireless LANs. In *Proceedings of ACM SIGCOMM '94*, London, UK, 1994.

[3] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination f unction. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, March 2000.

[4] M. Carvalho and J. J. Garcia-Luna-Aceves. A scalable model for channel access protocols in multihop ad hoc networks. In *Proceedings of ACM MobiCom '04*, Philadelphia, PA, September 2004.

[5] B. Chambers. The grid roofnet: a rooftop ad hoc wireless network. M.S. Thesis, MIT, May 2002.

[6] K. Chen, Y. Xue, and K. Nahrstedt. On setting TCP's congestion window limit in mobile ad hoc networks. In *Proceedings of IEEE International Conference on Communications*, Anchorage, Alaska, May 2003.

[7] Z. Fu, P. Zerfos, H. Luo, S. Lu, L. Zhang, and M. Gerla. The impact of multihop wireless channel on TCP throughput and loss. In *Proceedings of IEEE INFOCOM'03*, San Francisco, CA, April 2003.

[8] V. Gambiroza, P. Yuan, L. Balzano, Y. Liu, S. Shaefor, and E. Knightly. Design, analysis, and implementation of DVSR: a fair, high performance protocol for packet rings. *IEEE/ACM Transactions on Networking*, 12(1):85–102, February 2004.

[9] M. Gastpar and M. Vetterli. On the capacity of ad hoc wireless networks: the relay case. In *Proceedings of IEEE INFOCOM '02*, New York, NY, June 2002.

[10] M. Grossglauser and D. Tse. Mobility increases the capacity of ad hoc networks. In *Proceedings of IEEE INFOCOM '01*, Anchorage, Alaska, April 2001.

[11] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2), March 2000.

[12] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance anomaly of 802.11b. In *Proceedings of IEEE INFOCOM'03*, San Francisco, CA, April 2003.

[13] G. Holland, N. Vaidya, and P. Bahl. A rate-adaptive MAC protocol for multi-hop wireless networks. In *Proceedings of ACM MobiCom '01*, Rome, Italy, July 2001.

[14] X. Huang and B. Bensaou. On max-min fairness and scheduling in wireless ad-hoc networks: Analytical framework and implementation. In *Proceedings of ACM MobiHoc*, Long Beach, CA, October 2001.

[15] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu. Impact of interference on multi-hop wireless network performance. In *Proceedings of ACM MobiCom '03*, San Diego, CA, September 2003.

[16] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly. Ordered packet scheduling in wireless ad hoc networks: Mechanisms and performance analysis. In *Proceedings of ACM MobiHoc*, Lousanne, Switzerland, June 2002.

[17] R. Karrer, A. Sabharwal, and E. Knightly. Enabling large-scale wireless broadband: the case for TAPs. In *Proceedings of HotNets-II*, Cambridge, MA, November 2003.

[18] F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.

[19] A. Kherani and R. Shorey. Throughput analysis of TCP in multi-hop wireless networks with IEEE 802.11 MAC. In *Proceedings of IEEE Wireless Communications and Networking Conference*, Atlanta, GA, March 2004.

[20] J. Kurose and K. Ross. *Computer Networking: A Top-Down Approach Featuring the Internet*. Addison Wesley, 2nd edition, 2003.

[21] J. Li, C. Blake, D. S. J .D Couto, H. I. Lee, and R. Morris. Capacity of ad hoc networks. In *Proceedings of ACM MobiCom'01*, Rome, Italy, July 2001.

[22] H. Luo, J. Cheng, and S. Lu. Self-coordinating localized fair queueing in wireless ad hoc networks. *IEEE Transaction on Mobile Computing*, 3(1), January-March 2004.

[23] L. Massoulie and J. Roberts. Bandwidth sharing: objectives and algorithms. In *Proceedings of IEEE INFOCOM '99*, New York, NY, March 1999.

[24] T. Nandagopal, T. Kim, X. Gao, and V. Bharghavan. Achieving MAC layer fairness in wireless packet networks. In *Proceedings of ACM MobiCom'00*, Boston, MA, August 2000.

[25] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP Reno performance: A simple model and its empirical validation. *IEEE/ACM Transactions on Networking*, 8(2):133–145, April 2000.

[26] A. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM Transactions on Networking*, 1(3):344–357, June 1993.

[27] S. Pilosof, R. Ramjee, D. Raz, Y. Shavitt, and P. Sinha. Understanding TCP fairness over wireless LAN. In *Proceedings of IEEE INFOCOM'03*, San Francisco, CA, April 2003.

[28] B. Radunovic and J. Le Boudec. Rate performance objectives of multi-hop wireless networks. In *Proceedings of IEEE INFOCOM'04*, Hong Kong, March 2004.

[29] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly. Opportunistic media access for multirate ad hoc networks. In *Proceedings of ACM MobiCom '02*, Atlanta, GA, September 2002.

[30] M. Sanchez, T. Giles, and J. Zander. CSMA/CA with beam forming antennas in multi-hop packet radio. In *Proceedings of Swedish Workshop on Wireless Ad Hoc Networks*, March 2001.

[31] L. Tassiulas and S. Sarkar. Maxmin fair scheduling in wireless networks. In *Proceedings of IEEE INFOCOM '02*, New York, NY, June 2002.

[32] K. Xu, S. Bae, S. Lee, and M. Gerla. TCP behavior across multihop wireless networks and the wired internet. In *Proceedings of ACM International Workshop on Wireless Mobile Multimedia*, Atlanta, GA, 2002.

[33] K. Xu, M. Gerla, and S. Bae. How effective is the IEEE 802.11 RTS/CTS handshake in ad hoc networks? In *Proceedings of IEEE GLOBECOM'02*, November 2002.

[34] Y. Yang, J. Wang, and R. Kravets. Achievable bandwidth prediction in multihop wireless networks. Technical report, University of Illinois at Urbana-Champaign, Urbana, IL, 2003.

[35] Y. Yi and S. Shakkottai. Hop-by-hop congestion control over a wireless multi-hop network. In *Proceedings of IEEE INFOCOM'04*, Hong Kong, March 2004.

Appendix: TAP fairness reference model

Here, we present a formal definition that determines if a set of candidate allocated temporal shares (expressed as a matrix T) is TAP-fair. We define four objectives, namely *ingress aggregate, spatial bias, spatial reuse* and *temporal fairness*. For simplicity, we define TAP fairness for the case that all TAPs have equal weight; the definition can easily be generalized to include weighted fairness. Furthermore, we assume *fluid* arrivals and services in the idealized reference model, with all rates in the discussion below referring to instantaneous fluid rates. We refer to a *flow* as all uni-directional traffic between a certain ingress and egress pair, and we denote such traffic between ingress TAP$_i$ and egress TAP$_j$ as flow (i, j). Also, TA(i) denotes the aggregate of all flows (i, j) with ingress TAP$_i$. Associated to each flow is a number between 0 and 1 representing the fraction of time to be assigned to this flow from the ingress TAP to the next hop. Due to the flow preservation property, this number determines the fraction of time the flow is assigned at any other hop.

Consider a set of infinite-demand flows between pairs of a subset of nodes, with remaining pairs of nodes having no traffic between them. Denote T_{ij} as the candidate TAP fair share for flow (i, j). Let C_n be the capacity of link n, and let $t_n^{(i,j)}$ denote the time needed for flow (i, j) traffic to be transmitted on link n. Observe that if the capacity of the link from ingress TAP$_i$ to the next hop is C_i, we can write $T_{ij}C_i = t_n^{(i,j)}C_n$. Further, define contention neighborhood as a subset of the set of all links with the property that no two links from the subset can be active simultaneously, and there is no other link in the network such that by adding it to the subset, the property is preserved. The contention neighborhood defined as above contains both transmission and interference ranges. Note that a single link can belong to multiple contention neighborhoods. Denote L_n^k as the set of all links in link n's k^{th} contention neighborhood. The allotted time of link n's k^{th} contention neighborhood is then

$$\tau_n^k = \sum_i \sum_j \sum_{l \in L_n^k} t_l^{(i,j)}. \qquad (25)$$

Now we can write the following constraints on the matrix of allocated fair shares $T = \{T_{ij}\}$:

$$T_{ij} \geq 0, \text{ for all flows } (i, j) \qquad (26)$$

$$\tau_n^k \leq 1, \text{ for all } k \text{ and all links } n \qquad (27)$$

A matrix T satisfying these constraints is said to be feasible. Define the fair share of a TAP aggregate flow TA(i) and denote it as $T_a(i)$ such that

$$T_a(i) = \sum_j T_{i,j}. \tag{28}$$

DEFINITION 1. *A matrix of fair shares T is said to be TAP fair if it is feasible and if for each flow (i,j), T_{ij} cannot be increased while maintaining feasibility without decreasing $T_{i'j'}$ for some flow (i',j') for which*

$$T_{i'j'} \leq T_{ij}, \textit{when i=i'} \tag{29}$$
$$T_a(i') \leq T_a(i), \textit{ when } i \neq i'. \tag{30}$$

We distinguish two cases in Definition 1. First, in Equation (29), since flows (i,j) and (i',j') have the same ingress TAP, the inequality ensures fairness among a TA flow's sub-flows. In the second case, in Equation (30), flows (i,j) and (i',j') have different ingress TAPs. Thus, the inequality in Equation (30) ensures fairness among different TA flows.

Figure 21 illustrates the above definition. Assuming that all channel qualities are equal and all demands are infinite, the TAP fair shares are as follows: T_{14}=T_{13}=1/11, and T_{12}=T_{25}=T_{45}=2/11. If we consider flow (1,3), its fair share cannot be increased while maintaining feasibility without decreasing the fair shares of flow (1,4), or (1,3), where $T_{12} \geq T_{14}$, T_{13}, thus violating Equation (29). Finally, consider flow (4,5). Its fair share cannot be increased

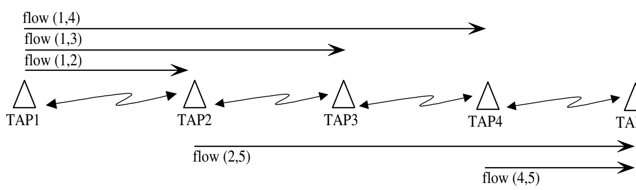


Figure 21: Illustration of TAP Fairness

while maintaining feasibility without decreasing the fair share of flow (1,4) or (2,5), and thereby violating Equation (30).

Below, we present an alternative way to determine if a set of allocated temporal shares is TAP-fair. To do so, we first define bottleneck link. Given a feasible matrix of fair shares T, we say that link n is a bottleneck link with respect to T for flow (i,j), and denote it by $B_n(i,j)$, if two conditions are satisfied. First, $\tau_n^k = 1$, i.e., there is no spare "time capacity". For the second condition, we distinguish two cases depending on the number of TAP-aggregated (TA) flows in the k^{th} contention neighborhood of link n. If TA(i) is not the only TA flow in the contention neighborhood, then $T_a(i) \geq T_a(i')$ for all TA flows TA(i'), and within TAP aggregate TA(i), $t_n^{(i,j)} \geq t_n^{(i,j')}$ for all flows (i,j') crossing link n. If TA(i) is the only TAP-aggregated flow in link n's k^{th} contention neighborhood then $t_n^{(i,j)} \geq t_n^{(i,j')}$ for all flows (i,j') crossing link n.

PROPOSITION 1. *A feasible fair share matrix T is TAP-fair if and only if each flow (i,j) has a bottleneck link with respect to T.*

Proof: Suppose T is TAP-fair. To prove the proposition by contradiction, assume that there exists a flow (i,j) with no bottleneck link. Then, for each link n crossed by flow (i,j) for which $\tau_n^k = 1$, there exists some flow $(i',j') \neq (i,j)$ such that one of Equations (29) and (30) is violated (which one depends on the relationship between flows (i',j') and (i,j)). Here, we present the proof for the case that Equation (30) is violated or more precisely when $T_a(i') > T_a(i)$. The proof is similar for the case when Equation (29) is violated. Now, we can write

$$\delta_n^k = \begin{cases} 1 - \tau_n^k, & \text{if } \tau_n^k < 1 \\ T_a(i') - T_a(i), & \text{if } \tau_n^k = 1 \end{cases} \tag{31}$$

where δ_n^k is positive. Therefore, by increasing the fair share of flow (i,j) by $\epsilon \leq \min\{\delta_n^k$: link n crossed by flow $(i,j)\}$ while decreasing by the same amount the fair share of the flow from TA(i') at contention neighborhood where $\tau_n^k = 1$, we maintain feasibility without decreasing the fair share of any flow TA(i') with $T_a(i') \leq T_a(i)$. This contradicts Definition 1.

For the second part of the proof, assume that each flow has a bottleneck with respect to T. To increase the fair share of flow (i,j) at its bottleneck link while maintaining feasibility, we must decrease the fair share of at least one flow from TA(i') (by definition we have $\tau_n^k = 1$ at the contention neighborhood of a bottleneck link), thus we must decrease $T_a(i')$. Furthermore, from the definition of bottleneck link, we also have that $T_a(i') \leq T_a(i)$. Thus, fair share matrix T satisfies the requirement for TAP fairness. ■

We make four observations about this definition. First, defining TAP-fairness in such a way, we are able to ensure all four objectives. Namely, by considering temporal shares we ensure temporal fairness, by satisfying the inequality in Equation (30) we ensure the ingress aggregate objective, by satisfying the equality in Equation (28) we ensure the spatial bias objective, and finally by allowing no spare "time capacity" we ensure spatial reuse. Second, we note that there can be multiple micro-flows with ingress TAP$_i$ and egress TAP$_j$. The fair share of such a micro-flow is T_{ij} divided by the number of micro-flows in flow (i,j). Third, observe that there are multiple policies to ensure the spatial bias objective. In Equation (28), we assumed a policy in which temporal shares from the ingress TAP to the next hop are considered. Another example policy is one in which temporal shares on the last link in the TAP network are considered. Finally, we note that Definition 1 is quite general. For example, observe that simply changing the objectives without changing the definition will result in a different matrix of fair shares T, and consequently a different fairness reference model. However, if we want to change the policy, we only need to change Equation (28). We demonstrate the differences with the example below.

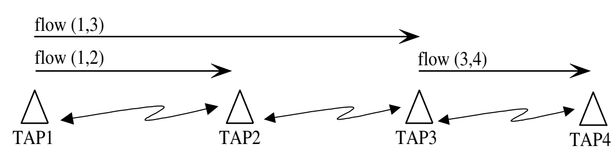


Figure 22: Illustration of Different Fairness Definitions

Consider the example in Figure 22, and three examples of the definition: one where objectives are defined as above, the second without the spatial bias objective, and the third without the ingress aggregate and spatial bias objectives. We want to determine $T = [T_{13}\ T_{12}\ T_{34}]$. Again assuming that all channel qualities are equal and all demands are infinite the appropriate fair share matrices are $T_1 = [1/5\ 1/5\ 2/5]$, $T_2 = [1/8\ 1/4\ 1/2]$, and $T_3 = [1/3\ 1/6\ 1/3]$.