Millimeter-Wave Blind Spots: Mitigating Deafness Collisions Using Frame Aggregation

Adrian Loch† Gek Hong Sim†‡
†Institute IMDEA Networks, Madrid, Spain
‡University Carlos III, Madrid, Spain

***Abstract*—Multi-gigabit-per-second data rates in millimeter-wave networks can result in an excessive amount of channel accesses. Transmitting a single packet at such rates often only requires a few microseconds. In contrast, the medium access control (MAC) overhead is about 20× larger. This is particularly harmful if one or more nodes are deaf due to the use of directional antennas. Deaf nodes cannot overhear ongoing transmissions, and thus cause a high number of collisions, exacerbating MAC overhead even more. In addition, the contention window of deaf nodes becomes very large since their transmissions often fail, resulting in unfairness. To mitigate this issue, we suggest increasing traffic burstiness deliberately to allow for higher frame aggregation at the physical layer. This reduces the number of medium accesses, and thus decreases the probability of collision in case of deafness. We adapt existing theoretical models to our scenario to show the effectiveness of this strategy. Moreover, we validate our scheme in a practical testbed operating in the 60 GHz band. In particular, our evaluation shows that our strategy improves fairness by 20% and throughput by 66% compared to a system that does not exploit deliberate burstiness.**

I. Introduction

Millimeter-wave networks can result in prohibitive channel access overhead. The key problem is that such networks provide multi-gigabit-per-second data rates while typical packet sizes are in the order of kilobytes. As a result, the physical layer transmission time of single packets is up to 20× smaller than the medium access control (MAC) overhead[1]. Additionally, millimeter-wave networks often suffer from deafness since nodes must use directional antennas to overcome the high attenuation at such frequencies. While this enables high spatial reuse [1], it is also prone to increase collisions. For instance, in Fig. 1, Station B is deaf to the communication between the Access Point (AP) and Station A since Station B is outside their beam angle. This often results in unfairness since Station B runs into a high number of retransmissions, resulting in an unnecessarily long backoff time. In contrast, Station A successfully transmits data and thus has a small contention window. Hence, Station A is more likely to win the contention.

The above combination of (a) very high MAC overhead and (b) deafness is highly harmful for millimeter-wave networks. Not only is the cost of accessing the channel prohibitive, but nodes are also prone to access it more often than in omni-directional communication due to frequent collisions. Existing standards for millimeter-wave networking such as 802.11ad [2] address overhead using frame aggregation. Essentially, transmitters aggregate multiple packets and access the channel only once to transmit all of them. Existing WiGig[2] hardware simply aggregates all packets available in the transmission queue. Related work optimizes the maximum aggregation size to, e.g., prevent the loss of large packets due to mobility [3].

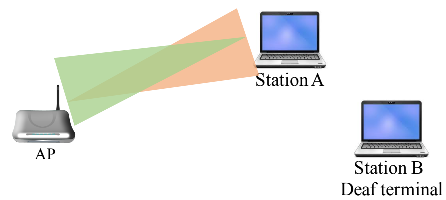


Fig. 1. Station B fails to overhear the transmission between Station A and the AP as it is not within the beam angle of Station A nor the AP.

In this paper, we exploit frame aggregation to reduce not just the MAC overhead but also the impact of deafness. In current implementations, a node continuously accesses the channel as long as it has at least one packet in its transmit queue. If it has more, it aggregates them, but otherwise it transmits even single packets. As a result, the medium is in use most of the time. If the node is in the "blind spot" of a second (deaf) node, the probability of collision is extremely high. Hence, we suggest allowing nodes to *wait* before transmission to form a burst of aggregated data. Introducing such deliberate traffic burstiness increases the idle time on the medium, and thus significantly reduces the probability of the deaf node colliding with the ongoing transmission. In particular, our contributions are as follows.

- We adapt existing theoretical models to include the increased medium idle time due to aggregation.
- We use the above models to show the benefits of our aggregation strategy for networks with deaf nodes.
- We validate our aggregation strategy in practice using commercial off-the-shelf WiGig hardware.

Our practical results show that our aggregation-based strategy improves fairness by 20% for a scenario with a deaf

[1]Transmitting a frame of size 1.5 KBytes on a 3.850 Gbps link requires only 3.1μs of the channel time while the MAC overhead is 60μs.

[2]WiGig is a precursor of the 802.11ad standard and operates similarly. We use WiGig hardware since 802.11ad devices are not available yet.

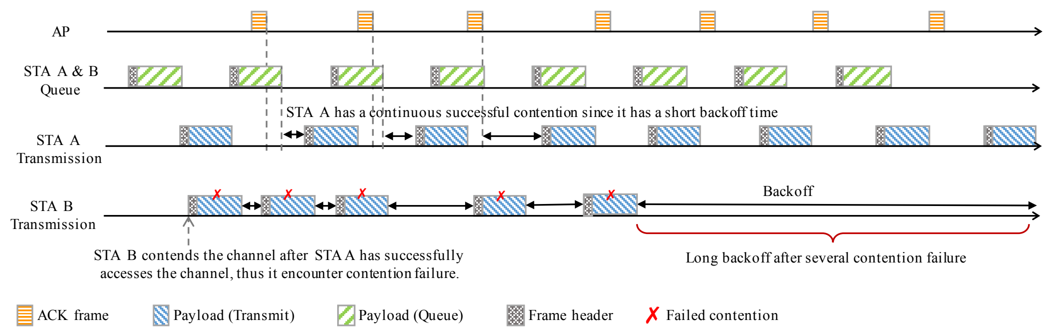


Fig. 2. Channel access frequency for bursty traffic with deaf stations.

node similar to Fig. 1. While our strategy cannot prevent deafness collisions entirely, it is simple to implement and fully local. In contrast, related work [4], [5] addressing deafness suggests solutions that are unsuitable for our case. For instance, the work in [4], and references therein, relies on omni-directional communication to broadcast control messages. Since such communication is either infeasible or unreliable in millimeter-wave networks, the applicability of this solution is questionable. Further, in [5] Singh et al. tailor a MAC protocol for 60 GHz communication and analyze it in theory. Their approach results in a time-division scheme that provides non-colliding scheduled access. While it addresses deafness in 60 GHz systems, it does not provide a solution for contention-based access.

The remainder of this paper is organized as follows. In Section II we explain the details of the deafness problem and how it leads to low fairness. After that, in Section III, we present our aggregation-based strategy and analyze its benefits in theory. Section IV then validates these benefits in practice. Finally, we discuss our evaluation results in Section V and conclude the paper in Section VI.

II. Problem Statement

In this section, we first give a brief overview on the contention-based access protocol for 60 GHz as specified in 802.11ad. Then, we present a detailed description of the existing problem of this channel access protocol.

A. Contention-based channel access in 802.11ad

Contention-based access in 802.11ad, which is also known as the contention-based access period (CBAP), mandates that non-transmitting and non-receiving stations (i.e., idle stations) listen to the channel in quasi-omnidirectional mode. This is particularly important since it allows other stations, regardless of their location, to request a transmission with such an idle station. Upon receiving a request from a transmitting station, the listening station switches to directional mode. To this end, it directs its beam towards the requesting station such that a transmission at a high data rate can be initiated. Once the data transmission is completed, the stations switch their beam pattern back to quasi-omnidirectional mode, and are thus ready to receive the next transmission request from another station.

B. Impact of deafness on fairness

Although listening in quasi-omnidirectional mode allows stations to potentially overhear transmission from any direction, a station is deaf towards a communication when it is not within the coverage range of the transmitter and/or receiver. The deafness problem is exacerbated when stations are communicating directionally. In directional communication, the beamwidth further limits the coverage range of the communicating pair. In particular, small beamwidths increase the severity of the deafness problem. Due to the high directionality of 60 GHz transmissions, the authors of [6] even considered 60 GHz links as pseudowired links.

Considering the scenario in Fig. 1, the deaf terminal (Station B) contends for the channel to transmit to the AP. Since the AP is directionally receiving from Station A, it misses the data transmission from Station B. Upon failure receiving an ACK from the AP, Station B increases its contention window and contends again for the channel. This procedure continues until Station B successfully communicates with the AP, which occurs when the AP switches back to quasi-omnidirectional mode. At this point in time, Station B also resets its contention window to a minimum value. However, if Station B continues to be unsuccessful, it suffers an unnecessarily long backoff time. This effect is illustrated in Fig. 2. Station A successfully accesses the channel multiple times, but Station B suffers from an extremely long backoff.

In a saturated scenario, stations attempt to access the channel immediately after a successful transmission, and after their backoff timer expires. The successful station, which is likely to have a smaller contention window, has a higher probability of successfully accessing the channel again. As a result, one station may occupy the channel for a longer time than the other, which impacts system fairness. Fig. 2 again shows this effect for our scenario in Fig. 1.

Additionally, 60 GHz systems such as 802.11ad achieve data rates of up to 7 Gbps. Hence, physical layer frames occupy an extremely short channel time. Due to this short transmission time, nodes need to contend for the channel much more often than in legacy non-multi-gbps systems in a certain time period. Such excessive channel access tremendously increases the collision rate between contending stations and/or the excessive backoff time due

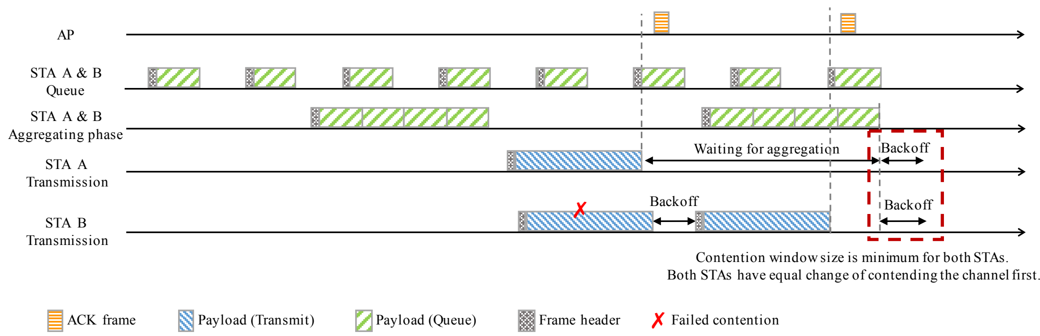


Fig. 3. Aggregation algorithm.

to contention failures, reducing the probability of a station successfully accessing the channel. This impact is much less pronounced in legacy wireless networks since transmission rates are significantly lower. For instance, a node transmitting packets of size 1518 bytes in legacy systems at 72 Mbps attempts to access the channel roughly 50× less often than a 60 GHz system operating at 3.850 Gbps.

III. Improving Fairness with Aggregation

This section presents our aggregation-based strategy to mitigate the excessive channel access and fairness problems described in Section II. Our strategy does not require any additional control messages, and is fully compatible with 802.11ad [2] and WiGig. In addition, we present a toy example that shows the effect of aggregation on deafness.

A. Intuition

Intuitively, aggregating traffic at a transmitting station reduces the channel access frequency and thus contention. For instance, consider a station which has five frames in its transmit queue. Without aggregation, the station has to contend five times for the channel. In contrast, it only needs to contend once for the channel if it uses aggregation. As a result, the probability of successfully allocating the channel time required to transmit the five frames increases exponentially. Further, collisions between stations become less frequent, and the overhead time due to, e.g., headers, inter-frame spacing times, and acknowledgments diminishes. Finally, this increases throughput since the ratio of payload with respect to the overhead time is greater.

Fig. 3 illustrates packet arrivals and contention phases for the example topology in Fig. 1. Before contending for the channel, stations A and B aggregate frames until reaching the maximum aggregation size, which in this example is four frames. Although Station B remains deaf to the transmission, it now has a wider window to contend for the channel. Precisely, it has a higher chance of accessing the channel since the channel is continuously free for a longer time, unlike the short idle time between transmissions in Fig. 2. The higher the aggregation, the longer the idle time.

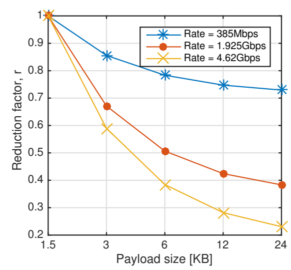


Fig. 4. Collision reduction factor r for different aggregation levels.

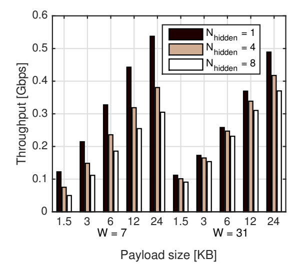


Fig. 5. Theoretical throughput results for increasing aggregation.

B. Toy example

Next, we evaluate our aggregation-based strategy in a theoretical toy scenario. To this end, we base our analysis on seminal work in the area of channel access in presence of hidden terminals. In particular, we adapt the models in [7] and references therein to our deafness scenario.

As discussed above, the number of channel accesses decreases when stations aggregate incoming packets before transmitting them. Thus, the probability of collision decreases as the aggregation level increases. Since this effect is not reflected in [7], we introduce a factor $r \in [0 \ldots 1]$ that takes into account the fraction of time during which the channel is idle due to aggregation. Fig. 4 shows the value of r for different frame payload sizes and data transmission rates. The former is directly related to the aggregation level. To compute the achievable throughput, we essentially multiply the probability of collision in the model of [7] by the reduction factor r. Using this adapted model, we obtain the results in Fig. 5, which confirm that the achievable throughput increases significantly with the payload size.

Fig. 5 shows the impact of the contention window size and of aggregation (i.e., payload size) on throughput. Each setting consists of one node which successfully communicates with an AP. Additionally, up to eight other nodes also try to transmit to the AP but are hidden from the first node. For simplicity, we assume a fixed contention window size. Hence, the transmission probability is $\tau = \frac{2}{W+1}$, where W is the contention window size. Further, the probability of collision is $p = 1-(1-\tau)^{n-1}$ where n is the total number of nodes. Since we assume that all nodes are deaf to each other in this example, the total expected time for successful packet collision (T_c) is equal to the time for transmission (T_s). For details on T_c and T_s, we refer the reader to [8].

In Fig. 5, we observe that increasing the payload size also increases the achievable throughput. However, the larger the minimum contention window size ($\mathrm{CW}_{\mathrm{min}}$), the longer is the idle time between transmissions, which in turn reduces the overall throughput. This is evident from Fig. 5 where the achievable throughput for high contention window sizes such as $W = 31$ is lower than that for $W = 7$.

IV. Evaluation

In this section, we first describe the experimental setup that we use to validate our strategy that improves fairness in 802.11ad networks with deaf nodes. After that, we present the performance metrics that we use in our experiments. Finally, we discuss our practical testbed results.

A. Experimental setup

Our experimental setup consists of two 60 GHz links, as shown in Fig. 1. However, 802.11ad access points are not yet available. Hence, we recreate an equivalent scenario using two Dell D5000 wireless docking stations. Each docking station connects to a Dell 6430U laptop, thus establishing a point-to-point connection. The D5000 implements the WiGig protocol, which operates quasi-identically to 802.11ad. This includes aggregation and contention-based access, allowing us to evaluate our strategy from Section III.

We place both laptops next to each other and at an angle of 45°, as shown in Fig. 6. They form "Link A" and "Link B" with both D5000 docking stations, respectively. This scenario recreates Fig. 1 — each D5000 acts as a station transmitting packets to an access point, and each of the laptops acts as a beam-steering direction of the access point. The directionality of the beam patterns of the Dell 6430U laptops is known to be limited [9]. As a result, the laptops overhear each other. This means that they cannot transmit simultaneously, similarly to an access point that can only transmit in one direction at a time. In contrast, the docking stations are deaf to each other. We connect them via gigabit Ethernet to two additional laptops, $\mathrm{L}_{\mathrm{TX}_\mathrm{A}}$ and $\mathrm{L}_{\mathrm{TX}_\mathrm{B}}$, that generate data streams with varying burstiness to induce different aggregation levels at the D5000s. Hence, in our scenario the D5000s transmit data to the 6430U laptops.

B. Metrics

We capture the performance of our system in terms of throughput, delay jitter, and fairness. We use Ostinato[3] to generate packets of size 1518 bytes at $\mathrm{L}_{\mathrm{TX}_\mathrm{A}}$ and $\mathrm{L}_{\mathrm{TX}_\mathrm{B}}$, and Wireshark[2] to capture them at $\mathrm{L}_{\mathrm{RX}_\mathrm{A}}$ and $\mathrm{L}_{\mathrm{RX}_\mathrm{B}}$. Based on the resulting traces, we can compute the throughput for any arbitrary time interval of an experiment. The delay jitter captures the periodicity of packet arrivals at the receivers. If only few collisions occur, the interval between packet arrivals is regular and thus the jitter is low. The more collisions occur, the larger the jitter is. Finally, we compute fairness using Jain's Fairness Index [10] as in Eq. 1,

$$F = \frac{\left(\sum_{i=1}^{n} \gamma_i\right)}{n \sum_{i=1}^{n} \gamma_i^2}, \quad (1)$$

where γ_i is the throughput of user i and n is the total number of users. Perfect fairness is achieved when the user's throughputs are equal resulting in $F = 1$, and perfect unfairness is when $F = 1/n$.

[3]Available at http://www.ostinato.org and https://www.wireshark.org/

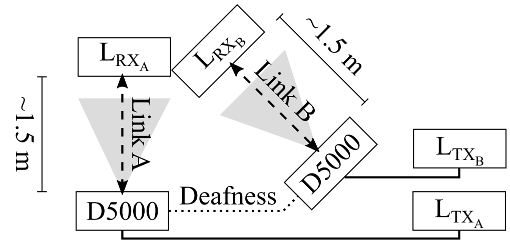


Fig. 6. Practical setup with deafness

C. Data streams

We use Ostinato to generate three different data streams. All streams result in the same net throughput but each of them exhibits a different burstiness. We set the throughput to 100 Mbps. While higher traffic loads are feasible, the D5000 limits us due to its automatic physical layer rate adjustment. We discuss this limitation in detail in Section V. Regarding burstiness, we set Ostinato to generate streams with one packet per burst, ten packets per burst, and one hundred packets per burst. The more packets in each burst, the higher the interval between bursts, since the throughput is constant. Such bursts induce aggregation at the D5000. Although we cannot adjust how much packets the D5000 aggregates in each transmission, earlier work [9] shows that the D5000 essentially aggregates as much packets as it has in its transmit queue, as long as this does not exceed the maximum aggregation defined in the 802.11ad standard. Hence, if Ostinato generates each packet individually, the D5000 is likely to contend for the medium for each single packet. In contrast, if we generate bursts of packets, the D5000 aggregates all of them. Our first experiment in Section IV-D validates this assumption. In other words, our three data streams induce three different aggregation levels.

D. Testbed results

This section presents and analyzes the results that we collected from the testbed we describe in Section IV-A.

1) Packet Aggregation: We first validate that increasing the number of packets per burst in Ostinato indeed (a) results in higher aggregation, and (b) reduces the medium usage time. To avoid interference from Link B, we operate Link A only. Further, we capture physical layer traces using a VubIQ 60 GHz downconverter [11] attached to an Agilent MSO-X 3034A oscilloscope. While the limited bandwidth of the oscilloscope hinders frame decoding, the traces allow us to determine the average aggregated frame length and the medium usage. Fig. 7 depicts our results. As expected, larger bursts yield longer frames at the physical layer due to increased aggregation. Most interestingly, we observe that the number of frames per second on the medium drops significantly with aggregation. Consequently, the probability of collision of packets of Link A with packets of Link B should drop significantly, too. In the subsequent experiments, we study this behavior.

2) Throughput: Next, we measure throughput operating both Links A and B simultaneously for increasing aggregation at the physical layer. Fig. 8 shows our results over a timespan of roughly 15 seconds with a resolution of 100 ms. We depict the throughput of both links side by side to allow for comparison of the general trends. However, the traces are not perfectly synchronous for both links since we cannot start the Wireshark trace at exactly the same moment in time on both L_{RX_A} and L_{RX_B}. That is, particular throughput variations on one link are not necessarily correlated to variations on the other link.

Still, our results clearly depict the impact of aggregation on the interference caused by deafness. For low aggregation, Fig. 8(a) shows that neither link can deliver the 100 Mbps traffic load that Ostinato generates. While Link A roughly achieves 80 Mbps, Link B suffers large oscillations at only half the average throughput. This unfair resource sharing is a direct result of deafness, as shown in Fig. 2 — Link A wins the contention much more often than Link B since its contention window is much smaller. As we increase the level of aggregation, resource sharing improves. In Fig. 8(b), we observe that Link A almost always achieves 100 Mbps. While Link B still suffers large oscillations, it also achieves 100 Mbps for a fraction of the time. Since the medium is empty more often (c.f. Fig. 7), collisions are less frequent despite deafness. This becomes particularly evident in Fig. 8(c). In this case, collisions are rare, allowing both links to continuously achieve 100 Mbps. That is, our strategy improves resource sharing and throughput significantly.

3) Delay Jitter: In our previous experiment, we claim that throughput degradation for low aggregation is due to collisions. To validate this, we study the delay jitter. The more collisions, the more irregular are the packet arrivals at

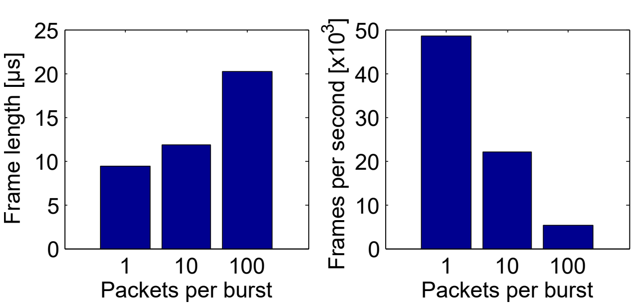


Fig. 7. Aggregation behavior with increasing burstiness

the receivers. In Figs. 9 and 10 we show the experimental Cumulative Distribution Function (CDF) of the interval between packet arrivals for Links A and B, respectively. This translates directly into the aforementioned delay jitter. For both links, we observe that no aggregation results in inter-packet arrival times which can be beyond one millisecond. This affects more than 60% of all packets on Link A, and more than 20% on Link B. If no collisions occurred, the inter-packet arrival time would be roughly constant, i.e., the CDF would be a quasi-vertical line. For increasing aggregation, the CDFs tend to this behavior.

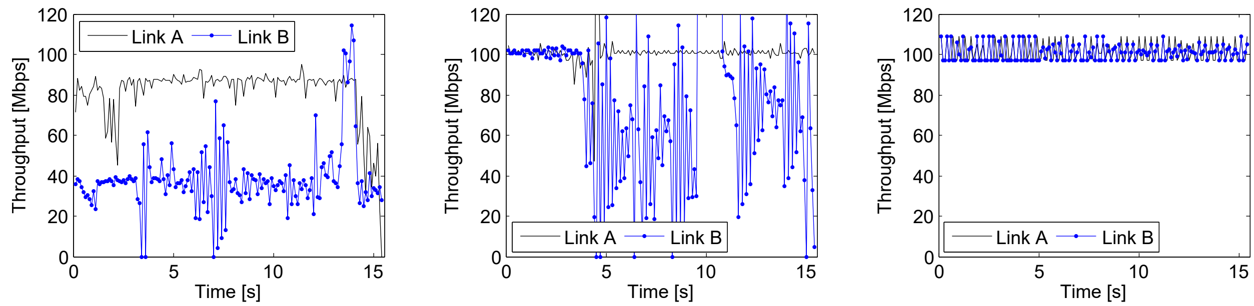
We conclude that no aggregation results in a significant number of collisions on the medium. Further the large variation of the packet arrival intervals has a huge impact on throughput — in comparison, the physical layer transmission of a single packet barely requires a few microseconds at 802.11ad rates. While more packets suffer such irregular arrival intervals on Link A than on Link B, most of them do not exceed 0.4 milliseconds. In contrast, on Link B a significant amount of packet intervals exceed one millisecond, ultimately resulting in lower average throughput.

4) Fairness: Finally, we study the average fairness $\bar{F}$, which we obtain by averaging the fairness of 100 experiments. Fig. 11 depicts our results. For all of the three aggregation levels, the average fairness tends to increase over time, as expected. Even if Link A wins the contention most of the times for the case with no aggregation, Link B also wins sometimes, thus slowly improving fairness. However, for higher levels of aggregation, we achieve perfect fairness (i.e., $\bar{F} = 1$) much faster. In other words, our aggregation-based strategy in Section III improves fairness in contention-based access with deaf nodes significantly.

V. Discussion

Our results show that high aggregation improves contention-based access in 802.11ad networks with deafness. However, to achieve high aggregation, transmitters must *wait* to receive enough packets. In our testbed results, we recreate this behavior at the application layer using Ostinato since we cannot control aggregation on the D5000. This raises the question on what the cost in terms of delay due to the aforementioned waiting is. This strongly depends on the rate at which the application generates data—the higher the rate, the less the station has to wait to form large bursts. A transmitter must decide whether this pays off for a particular data stream, according to its delay, throughput, and jitter requirements. However, designing such a decision mechanism is beyond the scope of this paper.

In our experiments, stations generate data at a rate of 100 Mbps. While higher rates are feasible, our experimental setup does not allow us to use them. The key problem is that the D5000 automatically switches to the lowest Modulation and Coding Scheme (MCS) in the event of collisions. For 802.11ad, this corresponds to a raw physical layer rate of 385 Mbps. As long as stations generate data at a lower rate, no queues build up and waiting to aggregate directly



(a) One packet/burst (b) Ten packets/burst (c) One hundred packets/burst

Fig. 8. Throughput of Links A and B for different aggregation levels

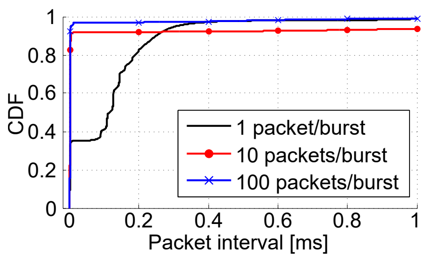


Fig. 9. CDF of packet arrival intervals on Link A

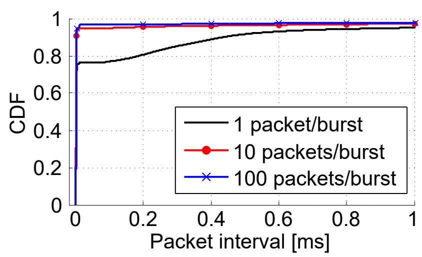


Fig. 10. CDF of packet arrival intervals on Link B

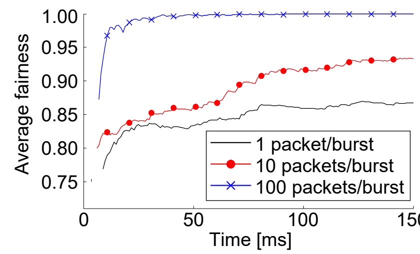


Fig. 11. Fairness among Links A and B

translates into a lower medium usage. At rates beyond 385 Mbps, queues quickly build up and thus no waiting is needed to aggregate. Still, this behavior of the D5000 is highly inefficient. A station should only switch to a lower MCS in case of actual channel degradation due to, e.g., increased attenuation, but not in case of transient collisions. For the latter, sticking to a high MCS ensures that the node can sustain a high throughput as soon as it wins the next contention. More sophisticated devices implementing such a policy would highly benefit from our aggregation-based strategy also in the case of multi-gbps rates.

VI. Conclusions

Directional communication undermines carrier sensing and introduces severe deafness problems. This leads to unfair channel sharing because deaf nodes collide more often and thus excessively increase their contention window. Moreover, MAC overhead in millimeter-wave networks is extremely high, and hence exacerbates even more the aforementioned unfairness. Our practical measurements using commercial off-the-shelf millimeter-wave hardware confirm that deafness has a significant impact on fairness. To address this issue, we propose a simple yet practical strategy based on frame aggregation. We deliberately cause traffic bursts to allow for higher aggregation at the physical layer. As a result, transmitters access the medium less often and incur less MAC overhead to transmit the same amount of data. This increases the medium idle time, hence significantly reducing the probability that a deaf node causes a collision. Our practical results on the aforementioned testbed show that our strategy improves throughput and fairness by up to 66% and 20%, respectively, compared to a system that does not cause deliberate burstiness. Future work includes implementing our strategy on software-defined radios to enable fine-grained aggregation control.

References

[1] X. Tie, K. Ramachandran, and R. Mahindra, "On 60 ghz wireless link performance in indoor environments," in PAM, Mar 2012.

[2] "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 3: Enhancements for very high throughput in the 60 GHz band," ISO/IEC/IEEE 8802-11:2012/Amd.3:2014(E), pp. 1–634, Mar 2014.

[3] S. Byeon, K. Yoon, O. Lee, S. Choi, W. Cho, and S. Oh, "Mofa: Mobility-aware frame aggregation in wi-fi," in ACM CoNEXT. ACM, Dec 2014.

[4] J. Qiao, L. Cai, X. Shen, and J. W. Mark, "Enabling multi-hop concurrent transmissions in 60 ghz wireless personal area networks," IEEE Transactions on Wireless Communications, vol. 10, no. 11, pp. 3824–3833, Nov 2011.

[5] S. Singh, R. Mudumbai, and U. Madhow, "Distributed coordination with deaf neighbors: Efficient medium access for 60 ghz mesh networks," in IEEE INFOCOM, Mar 2010.

[6] ——, "Interference analysis for highly directional 60-ghz mesh networks: The case for rethinking medium access control," IEEE/ACM Transaction on Networking, vol. 19, no. 5, pp. 1513–1527, Oct 2011.

[7] B. Jang and M. Sichitiu, "Ieee 802.11 saturation throughput analysis in the presence of hidden terminals," IEEE/ACM Transactions on Networking, vol. 20, no. 2, pp. 557–570, Apr 2012.

[8] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," IEEE Journal on Selected Areas in Communications, vol. 18, no. 3, pp. 535–547, Mar 2000.

[9] T. Nitsche, G. Bielsa, I. Tejado, A. Loch, and J. Widmer, "Boon and Bane of 60 GHz Networks: Practical Insights into Beamforming, Interference, and Frame Level Operation ," in ACM CoNEXT. ACM, Dec 2015.

[10] R. Jain, A. Durresi, and G. Babic, "Throughput fairness index: An explanation," Department of CIS, The Ohio State University, Tech. Rep., 1999.

[11] Vubiq. V60WGD03 60 GHz Waveguide Development System. [Online]. Available: http://www.pasternack.com/