Evaluating the Impact of Frame Aggregation on Video-Streaming over IEEE 802.11n Multihop Networks

Sascha Gübner and Christoph Lindemann
Department of Computer Science
University of Leipzig
Augustusplatz 10, 04109 Leipzig, Germany

{guebner, cl}@rvs.informatik.uni-leipzig.de

Abstract—**In this paper, we study the effects of frame aggregation on video streaming performance in a real-world IEEE 802.11n testbed. We quantitatively characterize delay, video quality, and mean aggregate size in two multi-hop scenarios. We observe that streaming applications naturally take advantage of frame aggregation, both in single- and multi-stream environments, and that a Full-HD video can be transmitted with excellent quality and delay, even on a 6-hop chain. Furthermore, we discover that limiting frame aggregation severely impacts both the delay and video quality. We believe these insights on frame aggregation help for developing an effective cross-layer design for video streaming over IEEE 802.11n multi-hop networks.**

Keywords: measurements in mesh networks, IEEE 802.11n, video streaming

I. INTRODUCTION

Mobile video is widely considered to become the next killer application for wireless networks [13]. According to a Cisco forecast [6], mobile video traffic will more than double every year till 2015 and has the highest growth rate of any application category measured within the forecast. This imposes high demands to the underlying cellular network infrastructure, which possibly can't keep up with this huge traffic increase. We suppose that multi-hop mesh networks may be an option to cope the increased traffic demands arising from mobile video streaming, especially from High Definition (HD) video. Other high potential fields of application for video streaming over mesh networks include real-time video streaming for emergency coordination or wireless video surveillance systems, e.g. for public safety or on building sites.

Although there are many promising scenarios, unfortunately, little is known about the behavior of video streaming in state-of-the-art mesh networks. To fully understand video streaming in new mesh networks, one must know insights on new features of the latest IEEE 802.11 standard, namely IEEE 802.11n [12]. With IEEE 802.11n, very high physical data rates up to 600 Mbit/s are achievable. However, these high data rates on the physical layer can only by harnessed at upper layers, if the medium access is efficient [17]. Therefore, IEEE 802.11n introduces frame aggregation on the MAC layer. With frame aggregation, multiple subframes, each with an own checksum, can be transmitted in an aggregated frame, with the overhead for medium access arising only once. On reception, the IEEE 802.11n MAC can correctly extract individual subframes even if other parts of the aggregated frame are erroneous. Frame aggregation will also be a key technology in future standards, like IEEE 802.11ac [11].

However, the influence of frame aggregation on video streaming performance in single- and multi-hop wireless networks is not sufficiently investigated. Therefore, we present in this paper the results of a measurement study on the impact of frame aggregation on video streaming performance in a real-world IEEE 802.11n mesh testbed. We herein focus on two typical scenarios. The first is a single video stream over a chain, e.g. when a mobile user streams video data from the Internet over wireless relay stations. In the other scenario, several sources stream and relay to one destination, as it occurs in wireless video surveillance. Our findings in these scenarios can be summarized as follows:

- Streaming applications naturally take advantage of frame aggregation, and the mean aggregate size increases both with video resolution and path length
- With IEEE 802.11n, a single Full-HD video can be transmitted with excellent delay and quality results, even over a 6-hop chain
- Limiting the frame aggregation severely impacts both the average delay and the quality of a video stream
- In a many-to-one scenario, up to 6 sources can stream an HD video to a common destination with acceptable quality, but favoring the nearest nodes
- In the same setup, the mean aggregate size increases almost linearly with each new stream
- Limiting the aggregation size in this scenario also severely impacts the video quality and average delay

We believe that these insights on frame aggregation are helpful to develop new enhancements for video streaming in IEEE 802.11n multi-hop networks.

The remainder of this paper is organized as follows. Section II summarizes related work on measurements in IEEE 802.11n networks and video streaming in wireless mesh networks. In Section III we introduce our experimentation environment; in Section IV the measurement results are presented. Finally, in Section V concluding remarks are given.

II. RELATED WORK

Cheng et al. [5] extensively studied the video streaming performance in an IEEE 802.11a multi-hop mesh testbed. They conclude that video has its special characteristics and that higher rates do not always imply higher quality. Dely et al. [8] studied the impact of IP packet aggregation and transmission opportunities on Voice over IP in an IEEE 802.11e testbed. Opposed to [5] and [8], we focus on the impact of frame aggregation in an IEEE 802.11n testbed.

Chan et al. [4] proposed a video-aware MAC rate adaptation scheme. It adapts frequency and timing of channel probing to video encoding rate and wireless channel variations. Jakubczak et al. [13] presented a cross-layer design between application and physical layer. They modified the signal transmission on the physical layer to be linearly related to a pixel's luminance. Opposed to [4] and [13], we focus on video streaming performance in an IEEE 802.11n multi-hop mesh network.

LaCurts et al. [15] analyzed traces gathered from different real-world wireless mesh networks deployed by Meraki, using both 802.11b/g and 802.11n devices. They mainly studied accuracy of SNR-based bit rate adaptation, the impact of opportunistic routing, and the prevalence of hidden terminals. Opposed to [15], we focus on the impact of frame aggregation on video streaming performance in an indoor mesh testbed.

In own previous work [10], we developed an analytical model for characterizing the effective throughput for multi-hop paths in IEEE 802.11n wireless mesh networks as a function of bit error rate, aggregation level, and path length. In [9], we studied the impact of aggregation and channel bonding on the achievable throughput on a multi-hop chain in our IEEE 802.11n mesh testbed. Opposed to [9] and [10], we focus in this paper on the impact of frame aggregation on video streaming performance.

In [3], Cai et al. presented an analytical model for studying the impact of frame aggregation and bidirectional transmission on voice and video performance for different aggregation schemes. In [16], Lee et al. detected through simulations that multiple-receiver frame aggregation can double the number of supported video streams. Li et al. [17] proposed an analytical model to derive the effective throughput and optimal frame and fragment sizes for transmissions with frame aggregation. Bononi et al. [2] proposed a joint channel assignment and multipath routing architecture for supporting robust video streaming over multi-radio mesh networks and showed its usability through simulations. Opposed to [2], [3], [16], and [17], we conducted measurements in a real-world IEEE 802.11n multi-hop testbed.

Pefkianakis et al. [18] studied MIMO based rate adaptation in 802.11n wireless networks and proposed a MIMO aware rate adaptation scheme. Pelechrinis et al. [19], [20] conducted experimental studies on the behavior of MIMO links in different topologies. They mainly focus on the impact of different 802.11n specific features on the peak performance [19] and packet delivery ratio under different physical data rates [20]. Opposed to [18], [19], and [20], we focus on video streaming performance in an IEEE 802.11n multi-hop mesh network instead of a 1-hop infrastructure mode WLAN.

Deek et al. [7] conducted experimental studies to examine the influence of channel bonding and interference on network performance, in terms of throughput and packet reception probability. Note, that frame aggregation was deactivated in most of their experiments, leading to reduced throughput compared to results in e.g. [9]. Shrivastava et al. [22] studied the impact of channel bonding and interference of 802.11g on 802.11n-links in a real testbed deployment. Opposed to [7] and [22], we consider video streaming performance in multi-hop IEEE 802.11n networks.

III. EXPERIMENTATION ENVIRONMENT

A. *Background on IEEE 802.11n*

The IEEE 802.11n standard features several enhancements for higher throughput in IEEE 802.11 wireless communication. Among others, the main features on the physical layer include channel bonding and the utilization of MIMO-specific features, like spatial division multiplexing or space-time block coding. Channel bonding allows an IEEE 802.11n transmitter to use two adjacent channels in parallel. While data rate increases, the communication becomes more sensible to interference, especially from legacy IEEE 802.11a/b/g links, thus collisions and erroneous transmissions occur more frequently. However, because we limited influence of external interference and to get insights on IEEE 802.11n's full potential, we activated this option.

The other MIMO-specific features are spatial division multiplexing and space-time block coding. Spatial division multiplexing increases the physical data rate by sending two or more independent data streams in parallel. Space-time block coding enhances the robustness of a transmission using a special coding. The choice, whether spatial division multiplexing or space-time block coding is utilized, is encoded in the modulation and coding scheme (MCS) of the transmitter. For our chipset, which is able to use 2x2 MIMO, MCS classes 0 to 7 use one data stream, while classes 8 to 15 use two data streams, encoded with spatial division multiplexing. In [9], we analyzed the distribution of utilized MCS classes and the limitation of MCS choice on throughput. Again, to study IEEE 802.11n's full potential, we didn't limit the MCS choice and let the bit rate adaptation algorithm choose the most appropriate MCS class.

Another important advancement of IEEE 802.11n concerns the MAC layer. IEEE 802.11n introduces frame aggregation to fully exploit the offered high physical data rates [17]. Frame aggregation, here MPDU aggregation, allows combining several MAC frames into one larger aggregated frame, with the overhead for medium access arising only once. Each correctly received frame can be acknowledged by the receiver in a block acknowledgement, reducing the number of frames to be retransmitted. Due to space limitations, we refer the interested reader to [12] Section 7.4a for details on MPDU aggregation. For our hardware, the maximum amount of frames that can be aggregated is limited to 32 frames.

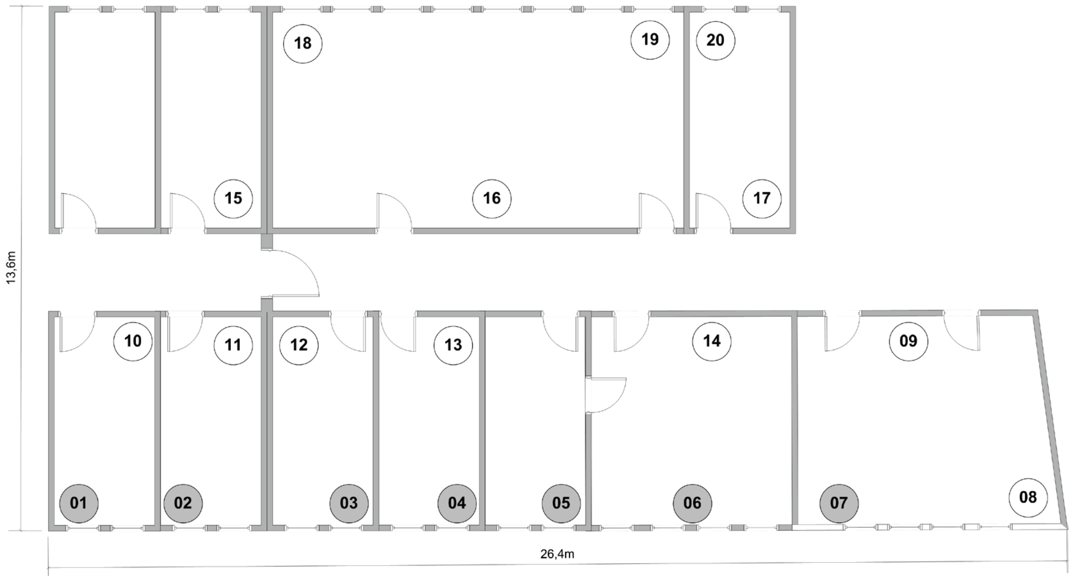

Figure 1. Indoor MIMO Mesh Testbed

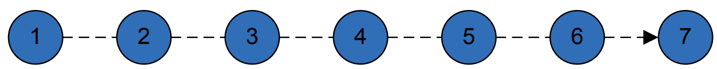

Figure 2. Scenario I: Chain topology (one stream with 6 hops)

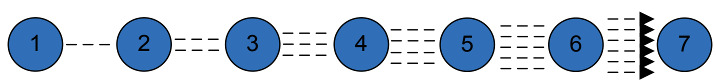

Figure 3. Scenario II: Many-to-one topology (six streams with *i* hops, *i*=1, 2, …, 6)

B. *Indoor MIMO Mesh Testbed Setup*

Our Indoor MIMO Mesh Testbed comprises 20 wireless mesh nodes located in 10 rooms in the department building, covering roughly 250 m². An overview of the testbed with the node locations is depicted in Figure 1. Each node consists of a Siemens ESPRIMO P2510 PC with an Intel Celeron 3.2 GHz processor, 80 GB HDD, 512 MB RAM, and a D-Link DWA-547 wireless PCI network interface card (NIC). This NIC is equipped with an AR 9223 Atheros chipset and three 5dBi omnidirectional antennas. The AR 9223 is able to support MIMO communication in the 2.4 GHz band. Each node runs openSUSE 11.2 as operating system with a modified kernel, based on version 2.6.34. Furthermore, we employ the ATH9K driver for Atheros chipsets.

Each node possesses a Gigabit Ethernet NIC to allow remote management of the nodes. Hence, wireless experiments can be managed from a remote computer, and traces can be copied and evaluated through the wired network. Table I shows a detailed description of hardware and software components of the testbed.

C. *Experimental setup*

For our experimental measurement study, we built up a 6-hop chain topology using the shaded nodes of our indoor MIMO mesh testbed depicted in Figure 1.

TABLE I. TESTBED OVERVIEW

Component	***Description***
PC	Siemens ESPRIMO P2510 Celeron 3.2 GHz, 512 MB RAM, 80 GB HDD
Wireless Card	D-Link DWA-547 PCI NIC equipped with 3 antennas
Chipset	Atheros AR 9223, operating at 2.4 GHz
Operating System	openSUSE 11.2 with kernel version 2.6.34

Note, that we limited the chain length to 6 hops, as we believe it's unlikely that longer path lengths will play a major role in a real-world deployment.

We placed the nodes to let the antennas face into the building to enrich the multi-path scattering, crucial for spatial division multiplexing. All nodes ran in ad-hoc mode, and we assigned static routing between them. To keep the nodes time-synchronized, we setup an NTP daemon on each node. Furthermore, we activated channel bonding to achieve the highest possible performance. As bit rate adaptation algorithm, we used *Minstrel HT*, the default rate adaptation algorithm for 802.11n in Linux. *Minstrel HT* is an advancement of the widely used SampleRate algorithm [1] by Bicket.

Because our devices use channel bonding and run in the 2.4 GHz band, we conduct the experiments at night or during the weekend. In these time slots little interference due to

802.11a/b/g background traffic occurred. We noticed in a previous long-term experiment that during the working hours, between 8am and 8pm, the measured goodput was influenced by external interference, especially due to students who access the Internet wirelessly through their IEEE 802.11 equipped laptops.

For our experiments, we used the video performance evaluation framework EvalVid [14]. As video test sequences, we extracted the first 60 seconds of the Full-HD videos "Elephants Dream" and "Big Buck Bunny" from [21]. Note, that although, due to space limitations, only the results of the first video are presented, we validated them also with the second one. We employed the open source encoder *x264* [23] with default settings to create single-layer video sequences, encoded in the widely used video coding standard H.264/MPEG-4 AVC. We encoded each video in different resolutions according to Table II. We set the Group-of-Picture size to equal the Frames-per-second, which are 24. The H.264 coding profile used was *High*. With these video samples, we conducted our streaming experiments starting from an initially idle system. We conducted 10 independent replicates of each experiment and derived the considered performance measures with 95% confidence level. The widths of the confidence intervals are depicted as bars in the plots.

For the actual streaming, we employed the EvalVid program *mp3trace*, which transmits a video from the source to a given destination, i.e. it uses unicast transmission. We used UDP as transport protocol to limit the influences of TCP's exponential backoff and retransmit mechanisms. We run the Linux tool *tcpdump* on the time-synchronized source and destination nodes, to trace delay and loss for each packet. With these traces, EvalVid is able to calculate delay and loss on a per-video-frame basis and can further reconstruct the video as it would be seen at the receiver. We use the sender and receiver side videos to calculate the Peak Signal-to-Noise Ratio (PSNR). PSNR is a standard objective metric to measure the quality of an encoded or transmitted video compared to the original one. Values around 40dB are considered as high quality video without any observable defect, while videos with a PSNR of 25dB to 30dB could still be acceptable, although visible artifacts exists [4].

We chose two scenarios for our experiments as examples for the application of video streaming. Figure 2 depicts Scenario I that considers a single video stream, transmitted over several hops on a chain topology. This is exemplary for e.g. live television streaming over a wireless mesh with an Internet gateway. However, Scenario II, depicted in Figure 3, considers several video streams in parallel. This is, for example, the case in wireless video surveillance, when several video cameras stream their data in a meshed manner to a central data collection and processing entity.

TABLE II. OVERVIEW OF USED VIDEO SAMPLES

Name	*Resolution*	*Coding Level*	*Average Rate [kbit/s]*
QVGA	320×240	1.3	315
VGA	640x480	3	960
HD	1280x720	3.1	2376
Full-HD	1920x1080	4	5510

IV. MEASUREMENT RESULTS

As an initial experiment, we stream all video samples over one hop with default configuration. We plot the sending rate for the 4 videos in Figure 4. We observe that the different videos share several rate peaks, occurring from fast changing and less compressible scenes, rising up to 13 Mbit/s for the Full-HD sample. In the next step, we gradually increase the path length up to 6 hops and calculated the average delay and PSNR of each video frame. All videos had excellent values; therefore, we neglect the corresponding curves. Instead, we concentrate on the most challenging video sample, the Full-HD one. We wanted to study the impact of frame aggregation on video streaming performance, in terms of delay and PSNR.

Therefore, we vary the maximum allowed number of MAC frames per aggregate. We denote this as the maximum aggregate size. We plot the results in Figure 5 and Figure 6. We can observe in both figures that decreasing the aggregation severely impacts the streaming quality. In Figure 5, we note that for all maximum aggregate sizes the delay rises almost linearly till a certain hop count, where it suddenly grows faster. For maximum aggregate size 2, this is already at 3 hops, while it's at 4 hops for maximum aggregate size 4 and 5 hops for larger sizes. We further note that the increase is slighter for larger maximum aggregate sizes. This effect is due to the increased load and medium contention when, on the one hand, path length increases and, on the other hand, maximum aggregate size decreases. The latter causes more small aggregates to be transmitted, thus increasing the overall time when the medium is blocked. In Figure 6, we observe the same effect. With decreasing maximum aggregate size, the achievable effective capacity of the medium decreases too. This leads to higher loss probabilities and, therefore, a lower PSNR. We further note that for a maximum aggregate size of 32 a Full-HD video can be transmitted without noteworthy distortion, even over a 6 hop chain. Note, that we neglected the confidence intervals here for clarity, as they overlapped confusingly.

Now we want to have a closer look at the aggregation itself. Figure 7 depicts the mean aggregate size on the last hop, i.e. the 6th node in Figure 2. We notice that with increasing resolution also the mean aggregate size rises.

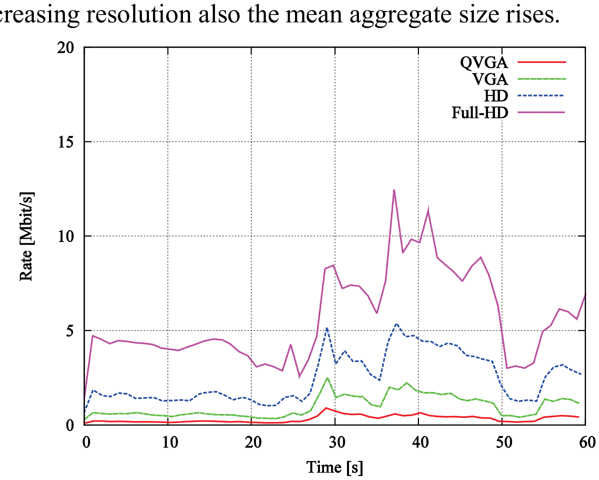

Figure 4. Sending rate vs. time for different resolutions in Scenario I

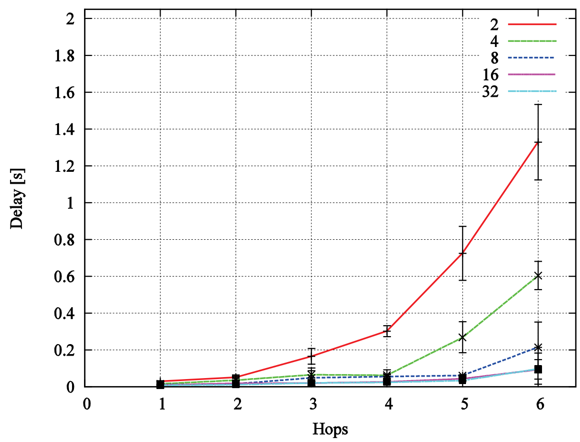

Figure 5. Average end-to-end delay vs. hops of a Full-HD video for different maximum allowed aggregate sizes in Scenario I

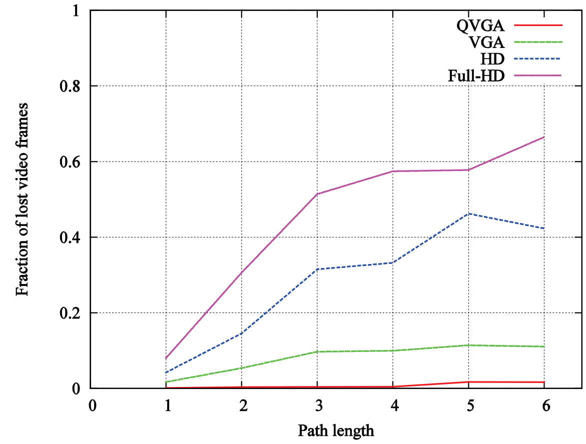

Figure 8. Fraction of lost video frames vs. path length on a 6-hop chain for different resolutions in Scenario II

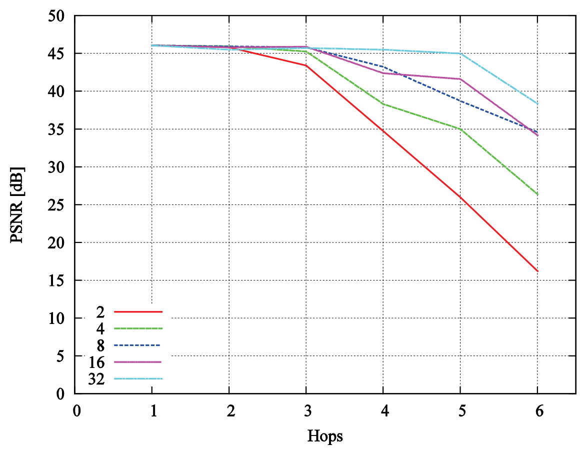

Figure 6. Average PSNR vs. hops of a Full-HD video for different maximum allowed aggregate sizes in Scenario I

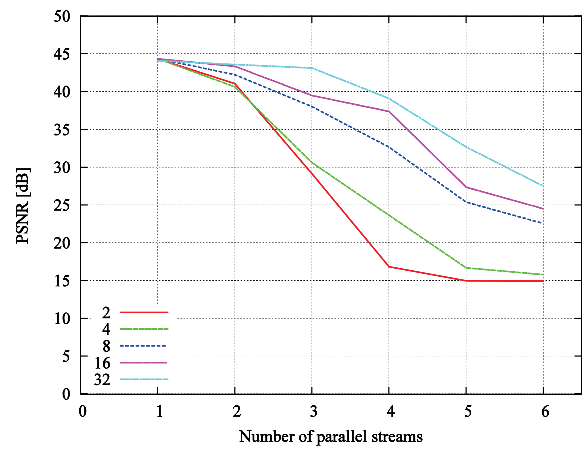

Figure 9. Average PSNR vs. number of parallel streams of an HD video for different maximum allowed aggregate sizes in Scenario II

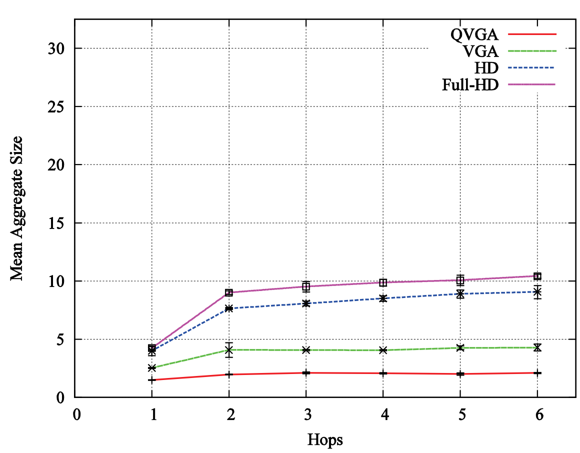

Figure 7. Mean aggregate size on last hop vs. hops for different resolutions in Scenario I

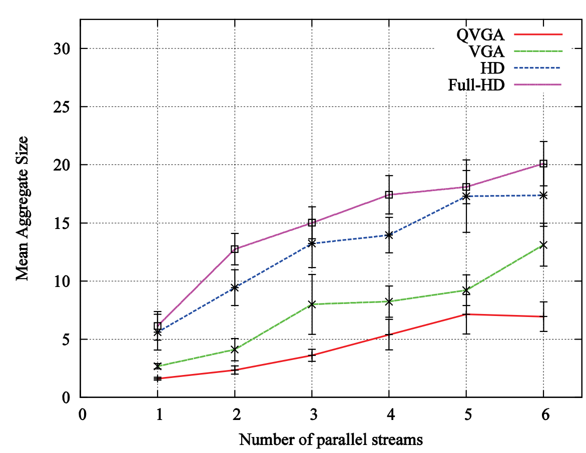

Figure 10. Mean aggregate size on last hop vs. number of parallel streams for different resolutions in Scenario II

This is evidence that frame aggregation automatically adapts to higher traffic load, i.e. with higher medium contention, the packet queue increases leading to higher aggregation opportunities. This is also why the mean aggregate size increases slightly with the path length.

The next experiment is conducted for Scenario II. Here, we stream several videos in parallel, i.e. each node in the chain streams a video to the destination node. Note, that we start the video sequences with some delay to avoid cumulating peaks in the sending rate. This could be the case, if several surveillance cameras in a row consecutively record a moving object. In Figure 8, we first have a closer look on the loss each video stream suffers in the full 6-hop scenario, i.e. 6 nodes stream in parallel to the destination node. We observe that video streams from nodes farther to the destination suffer higher losses. For the Full-HD sample, these losses become rapidly rising over 50%. We suppose this is due to queue drops at relaying nodes that can't access the medium quick enough.

Now we look at the impact of frame aggregation on the average Peak Signal-to-Noise Ratio, depicted in Figure 9. We observe that it generally decreases linearly with the number of parallel video streams, like in Figure 6. Note, that the delay curve, not shown here, is comparable to Figure 5. We left it due to space limitations. Furthermore, we notice in Figure 9 that, with full frame aggregation enabled, we can stream several HD-videos in parallel with acceptable video quality, even over 6 hops.

In Figure 10, we again vary the chain length but look now closer on the mean aggregate size on the last hop. We observe that the mean aggregate size almost proportionally increases with the number of video streams that traverse the node. This indicates that frame aggregation helps to alleviate the increasing traffic load, by enlarging the transmitted physical data units, thus minimizing the time, when the medium is blocked.

V. Conclusion

We analyzed the impact of IEEE 802.11n frame aggregation on video streaming performance, in terms of delay and video quality. We further revealed details on the dependency of number of video streams, video resolution, and path length on the aggregation level. We showed that a Full-HD video can be transmitted with excellent delay and quality, even over a 6-hop chain, and that IEEE 802.11n is also able to support several HD videos in a many-to-one scenario. We further showed that limiting the frame aggregation severely impacts both the average delay and the quality of a video stream, in both considered scenarios.

For future work, we will use our findings to optimize the alignment of video transmission and frame aggregation to increase the number of supported videos with 802.11n.

References

[1] J. Bicket, Bit-rate Selection in Wireless Networks, Masters Thesis, MIT, Cambridge, USA, 2005.

[2] L. Bononi, M. Di Felice, A. Molinaro, and S. Pizzi, A Cross-Layer Architecture for Robust Video Streaming over Multi-Radio Multi-Channel Wireless Mesh Networks, *Proc. ACM MobiWac*, Tenerife, Canary Islands, 2009.

[3] L. Cai, X. Ling, X. Shen, J. W. Mark, and L. Cai, Supporting voice and video applications over IEEE 802.11n WLANs, *Springer Wireless Networks*, **15**, 2009.

[4] A. Chan, H. Lundgren, and T. Salonidis, Video-Aware Rate Adaptation for MIMO WLANs, *Proc. IEEE ICNP*, Vancouver, Canada, 2011.

[5] X. Cheng, P. Mohapatra, S.-J. Lee, and S. Banerjee, Performance evaluation of video streaming in multihop wireless mesh networks, *Proc. ACM NOSSDAV*, Braunschweig, Germany, 2008.

[6] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010–2015, http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html, 2011.

[7] L. Deek, E. Garcia-Villegas, E. Belding, S.-J. Lee, and K. Almeroth, The impact of channel bonding on 802.11n network management, *Proc. ACM CoNEXT*, Tokyo, Japan, 2011.

[8] P. Dely, A. Kassler, N. Bayer, and D. Sivchenko, An Experimental Comparison of Burst Packet Transmission Schemes in IEEE 802.11-based Wireless Mesh Networks, *Proc. IEEE GLOBECOM*, Miama, USA, 2010.

[9] J. Friedrich, S. Frohn, S. Gübner, and C. Lindemann, Understanding IEEE 802.11n Multi-hop Communication in Wireless Networks, *Proc. IEEE WinMee*, Princeton, USA, 2011.

[10] S. Frohn, S. Gübner, and C. Lindemann, Analyzing the Effective Throughput in Multi-Hop IEEE 802.11n Networks, *Elsevier Computer Communications*, **34**, 2011.

[11] IEEE P802.11 - TASK GROUP AC, Specification Framework for TGac, IEEE, 2011.

[12] IEEE 802.11n: Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput, IEEE, 2009.

[13] S. Jakubczak and D. Katabi, A cross-layer design for scalable mobile video, *Proc. ACM MOBICOM*, Las Vegas, USA, 2011.

[14] J. Klaue, B. Rathke, and A. Wolisz, EvalVid - A Framework for Video Transmission and Quality Evaluation, *Proc. TOOLS*, Urbana, USA, 2003.

[15] K. LaCurts and H. Balakrishnan, Measurement and Analysis of Real-World 802.11 Mesh Networks, *Proc. ACM IMC*, Melbourne, Australia, 2010.

[16] K. Lee, S. Yun, and H. Kim, Boosting Video Capacity of IEEE 802.11n through Multiple Receiver Frame Aggregation, *Proc. IEEE VTC*, Marina Bay, Singapore, 2008

[17] T. Li, Q. Ni, D. Malone, D. Leith, Y. Xiao, and R. Turletti, Aggregation with Fragment Retransmission for Very High-Speed WLANs, *IEEE/ACM Transactions on Networking*, **17**, 2009.

[18] I. Pefkianakis, Y. Hu, S. H. Wong, H. Yang, and S. Lu, MIMO Rate Adaptation in 802.11n Wireless Networks, *Proc. ACM MOBICOM*, Chicago, USA, 2010.

[19] K. Pelechrinis, T. Salonidis, H. Lundgren, and N. Vaidya, Analyzing 802.11n Performance Gains, *Proc. ACM MOBICOM* (poster session), Beijing, China, 2009.

[20] K. Pelechrinis, T. Salonidis, H. Lundgren, and N. Vaidya, Experimental characterization of 802.11n link quality at high rates, *Proc. ACM WINTECH*, Chicago, USA, 2010.

[21] P. Seeling and M. Reisslein, Video Transport Evaluation With H.264 Video Traces, *IEEE Communications Surveys & Tutorials*, to appear

[22] V. Shrivastava, S. Rayanchu, J. Yoon, and S. Banerjee, 802.11n Under the Microscope, *Proc. ACM IMC*, Vouliagmeni, Greece, 2008.

[23] X264, H.264/AVC Encoder, http://www.videolan.org/developers/x264.html