Project 3: small HTTP packets

Here are some HTML-only websites (as of March 202). The numbers represent their rank at one point in a ranking of top websites. The data originally came from whynohttps.com.

11. baidu.com
61 . xinhuanet.com
80. apache.org
102. babytree.com
123. tianya.cn
129. go.com
153. gnu.org
160. soso.com
166. china.com.cn
280. drudgereport.com
295. nginx.org
341 . washington.edu
348. thestartmagazine.com
365. rlcdn.com
477. chinadaily.com.cn
494. yimg.com
525. gmw.cn
526. eastday.com
537. eepurl.com

I loaded each site with Chrome, and recorded all the traffic at my router, 23,962 packets in all. That file is **project3.pcap**. This file can be opened with Wireshark to see the packets.

Both data and ACK packets are shown. Generally speaking, I would expect that, for HTTP connections, almost all the data would be in the downstream packets (to the client), and the upstream packets (from the client) would be ACKs only.

But now let's use the Wireshark statistics => packet lengths option. That's all packets in either direction, so let's just look at packets with destination equal to my router, which is 92. 68. . 0. I do that by entering a Wireshark **display filter** into the box at the bottom. While I'm at it, I restrict attention to tcp packets involving port 80 at either end (I can specify tcp.srcport if I wanted the source port):

ip.dst == 92. 68. . 0 and tcp.port == 80

This is what we get:

Packet Lengths:

Topic / Item	Count	Average	Min val	Max val	Rate (ms)	Percent	Burst rate	Burst start
Packet Lengths	25490	1310.07	62	1468	0.0911	100%	1.1600	20.461
0-19	0	-	-	-	0.0000	0.00%	-	-
20-39	0	-	-	-	0.0000	0.00%	-	-
40-79	**1947**	**66.74**	**62**	**78**	**0.0070**	**7.64%**	**0.4500**	**24.556**
80-159	180	124.47	80	159	0.0006	0.71%	0.1800	29.609
160-319	169	235.17	160	317	0.0006	0.66%	0.0500	10.894
320-639	440	476.20	320	638	0.0016	1.73%	0.1000	58.891
640-1279	516	927.21	641	1275	0.0018	2.02%	0.0800	28.895
1280-2559	22238	1462.07	1280	1468	0.0795	87.24%	0.9800	19.605
2560-5119	0	-	-	-	0.0000	0.00%	-	-
5120 and greater	0	-	-	-	0.0000	0.00%	-	-

Now we just have 25,490 packets. 87% of them are in the size group 1280-2559 (for all intents and purposes, this is 1280-1514; packets cannot be larger than that). This is what I expect.

But 7.64% are in the size range 40-79. **What are these small packets**? ACKs? If so, what data is being sent from the client?

Your assignment is to try to figure out what all these smaller packets are for. Use as many of Wireshark's features as possible. Because this is http (not https) traffic, you can observe the actual data.

You can view just these packets by adding the following to the previous display filter (joined with **and**):

and frame.len >= 40 and frame.len < 80

Here are some useful tools (aside from the creative use of display filters):

- Analyze => Conversation Filter => TCP
- Statistics => Conversations
- Analyze => Follow => TCP Stream
- Local port numbers and remote IP addresses determine connections

One approach is to try ~10 packets at random, and see what kind of traffic flow they are part of.

To submit, write up a report discussing your analysis.