# Appendix A

# Congestion Avoidance and Selective Retransmission Policies for TCP

Our TCP SACK-NewReno implementation obeys standard congestion avoidance policies and rules for selective acknowledgments (SACKs) as specified in [129] and [83], with the following extensions.[1] The following extensions apply whether or not SACK is enabled for a given connection:[2]

1. Initialize a new state variable, *snd_recover*, to the value of *snd_una* upon connection start.
2. Upon receiving three duplicate acknowledgments, if the sequence number acknowledged is greater than or equal to *snd_recover*, then set *snd_recover* equal to *snd_max*, and perform fast retransmit according to [129].
3. If, while in fast recovery phase, a segment acknowledging new data is received and the sequence number acknowledged is greater than or equal to *snd_recover*, then exit fast recovery by setting *snd_cwnd* to either *snd_ssthresh* or the amount of outstanding data in the network plus one segment, whichever is smaller.
4. While in fast recovery phase, if a segment acknowledging new data is received, and the sequence number acknowledged is less than *snd_recover*, if SACK is not enabled for the connection then retransmit the next unacknowledged segment. Additionally, whether or not SACK is enabled, partially deflate the (inflated) *snd_cwnd* by the amount of new data acknowledged, add back one segment to *snd_cwnd*, and call *tcp_output()*.

In addition, if SACK is enabled for a given connection, the following rules apply to retransmissions and new data transmissions during the recovery phase:

5. A given segment is considered "eligible" for retransmission if it has not already been retransmitted and if either three duplicate acknowledgments have arrived for the segment just prior to it or the SACK information implies that the receiver is holding a segment that was sent at least three segments beyond the given segment.

[1] This description assumes a TCP implementation similar in structure to Berkeley-derived TCP implementations.

[2] These first four guidelines are for the TCP NewReno portion of the implementation and have been accepted as (experimental) RFC 2582 within the IETF [42].

6. While in fast recovery, upon reception of each ACK that does not end the fast recovery phase, the TCP sender first checks whether there are any eligible retransmissions to be sent. If so, one such retransmission is sent. If not, the TCP sender inflates *snd_cwnd* by one segment and attemps to send one or more new segments if permitted by the window.

7. When *snd_max* is greater than *snd_nxt* (e.g., following a TCP timeout), any SACK information received subsequent to the timeout is used to avoid retransmitting data for which the receiver is sending a SACK.