3.3, middlebox behavior, you blame inability to detect SYN/ACK option stripping on to the statelessness of your tool, but that doesn't seem to be accurate.

4.3, "We hope that these middleboxes are aware of SACK"... are they or are they not? Hoping doesn't do...

4.7, the main trouble is that simply duplicating options onto all segments, as TSO does isn't enough in many cases...

Reviewer #4

Strengths: Detailed catalog of behavior "in-the-wild". Relatively well thought out experiments, and some unexpected findings.

Weaknesses: Writing is repetitive. A number of 'problems' were due to mis/aggresively configured HTTP proxies and do not reflect behavior of other TCP flows.

Comments to Authors: Is there a reason for the distribution of the 142 networks in the paper? They are not representative of Internet traffic volume, and it is not clear that they are representative of middlebox behavior.

The related work should refer to the TCP Sidecar paper (IMC 2006) which describes how active measurements can go through middleboxes.

For the most part, the results are not surprising. Application layer gateways try to parse data and fail if they don't see everything. Proxies/middleboxes that regenerate sequence numbers don't preserve options that refer to literal sequence numbers (much like FTP PORT command and NAT interaction). However, the paper is valuable in that it provides a systematic catalog of anecdotal behavior.

The paper is repetitive, and the writing verbose. The information here can be fit into a very good seven page paper, as opposed to the loose fourteen pager that is presented. Section 4.5+ adds very litle that has not already been stated or could be put inline with the other results. The entire issue with the seven HTTP proxies is a red herring and could be mentioned once in a footnote.

The paper is also very difficult to follow since the tables are scattered all over, making it difficult to refer to them while reading the text.

Reviewer #5

Strengths: An extensive study. I particularly like that the test traffic is controlled at both ends of the path (by having clients voluntarily download and run a test program) as it is much less limiting than client or server only approaches. Useful results that are not otherwise known. Anyone seriously looking to extend TCP will want to read this paper.

Weaknesses: The paper is a bit ad hoc. The questions that are asked about how middleboxes handle TCP are mostly driven by efforts to deploy a multipath TCP, which is both good (they are relevant questions!) and bad (as it causes them to focus heavily on sequence numbers and is unclear that they will cover the needs of other extensions that may come). The paper would benefit from being a bit more systematic in its exploration of the space.

Comments to Authors: Thanks for an interesting paper; I have relatively few comments.

I think your paper will benefit from stepping back a bit to separate it from MTCP. What other aspects of middlebox behavior might be important for extensions? For example, are there games with flow control? What about the window scale, MSS, and authentication options, etc., as they at least seem worth some study? One exercise you might attempt is to go through all known extensions and make a a table with the TCP header fields or other properties/invariants on which they depend for correct behavior.

People are likely to read your paper to get guidance on what is safe/unsafe. Thus you might provide an easily accessible and complete summary of the takeaways (that is more comprehensive and standalone than in the conclusion).

I'd also be interested in recommendations to middlebox developers for what to do or not to do wrt unsupported options to maximize the ability for future extensions, i.e., how can we make the future better. Is this already done?

Section 5 seems misnamed. It is really a set of three case studies of how TCP extensions should work with TCP options. The bit that I found odd here is that the extensions have been designed in light of what was known about middlebox behavior by the authors. The presentation makes it sound like there are these new protocols that are to be assessed to see how they will interact with middleboxes -- and they are found to be mostly compatible.

Response from the Authors

First, we made the description on our dataset clearer; for example, how we collect data, and how we identified the venue.
Second, we added a few paragraphs guiding middlebox design that will work together future TCP extensions.
Third, we added a sentence describing how HTTP proxies behave on manually verified two paths.
In addition to these update, we've polished the entire document to provide more precise and explicit information. Also, we added results of tests for Large Receive Offload (LRO) as supplemental information.