Gmail +talk BETA

Stefan Karpinski <stefan.karpinski@gmail.com>

Paper comments - part 1

Kimaya Mittal <kimaya@gmail.com> **Wed, Dec 13, 2006 at 2:40 PM**
To: stefan.karpinski@gmail.com, sgk@cs.ucsb.edu

Hey Stefan,

I read your paper till section IIIC. Here are my comments. I'm not sure if I'll have the time to read the rest before leaving for Hawaii tomorrow morning, but will certainly try.

Overall, the paper looks great. I don't really have any comments about the technical content, just some suggestions about organization/emphasis, and some questions that reviewers might have and that you might want to address.

1. Abstract: The way the abstract reads currently, it seems to emphasize your evaluation and the finding that CBR and other synthetic models are highly unrealistic. Some reviewers could argue that this is showing the obvious, something that everyone already knows. Further, if you already know something is extremely inaccurate, quantifying the inaccuracy (factor of two or four or six) is not super interesting. I would rather shift the focus to your two other important contributions: (1) the statistical methodology and tools (definition of 'sufficiently realistic', measure of error, tests of statistical equivalence) - emphasize that these create a basic framework that is applicable to any new traffic model created in future - CBR is just an example you use to demonstrate the methodology; and (2) creating guidelines for developing more realistic traffic models. I think these two contributions make a much stronger impact than the evaluation part.

2. Introduction, para 3: You mention that a usage model is sufficiently realistic if it produces performance results that are statistically equivalent to those produced by real data. It made me wonder: would results produced by two different sets of real data (at least from similar network deployments) be statistically equivalent? If yes, can you motivate/demonstrate that? If not, then what does 'sufficiently realistic' really mean? (Later while reading section IIIA, it occured to me that maybe you could use two 24-hour traces and compare the statistical equivalence of results between them. Would that make sense?)

3. I like how the three contributions are enumerated towards the end of the introduction section. Maybe you could do that in the abstract as well?

4. When you are measuring if something is 'sufficiently realistic', how much impact does lower layer modeling have? Could something be sufficiently realistic w.r.t. a traffic trace under one set of lower layer models but not under another?

5. I would reorganize section III. Following your enumeration of contributions in the introduction, sections C and D pertain to contribution #1, while sections A and B pertain to contribution #2. I would present C and D first, and then present A and B as an application of the proposed methodology to evaluate the specific models (CBR and others).

6. I really like the beginning of section III.

7. Section IIIB: Minor comment: When listing TTU, TUU, etc. in the text, would be good to state that the expansions are in the table.

8. Equation (5) didn't make sense to me.

That's all for now. Sorry about the wordy comments - I don't have the time to make them more concise. Hopefully they make sense and you get what I'm trying to say. Mail back if you need clarification on anything - I will be checking mail until late tonight.

Good luck!
Kimaya