Internet Simulations:
Issues in Defining the Model

Sally Floyd
Network Research Group, Lawrence Berkeley National Laboratory
http://www-nrg.ee.lbl.gov/floyd/

DIMACS Workshop on End-to-End Network Modeling and Simulations

October 23-24, 1997

Credits are due to many other people, most notably to Deborah Estrin...

Overview of the talk:

- What are the problems in defining models in Internet research/simulations?
- Examples from my own research.
- Interspersed suggestions of (and requests for) tools that would help.

Defining the model: what is the problem?

- What level of detail or abstraction is appropriate?

- How to model a rapidly-changing world (that is changing in ways that are hard to predict)?

- What to model in a world of great heterogeneity (of underlying technologies, link bandwidth and propogation delay, levels of congestion, protocols, router scheduling and queue management mechanisms, etc.)

- How to handle issues of scale?

- Sensitivity to small changes in parameters of the model.

Propositions:

• Choosing the appropriate model is one of the key steps in a network research project.
Corollary: One of the criteria for evaluating work is the appropriateness of the problem definition and of the models chosen.
Corollary: The relationship between defining the model and simulating/analysis is two-way.

• Better tools can make this easier: richer tools and models for topology generation and traffic generation, tools for visualization of research-in-progress, tools for more explicitly varying the level of detail or abstraction, discussion in the network research community, etc.

• There are plenty of challenging and important questions of scale to address in network research. These will stress our modeling and simulation tools. (There are presumably also challenging and important questions that are not about issues of scale?)

Level of detail: TCP's retransmit timers

- The issue:

The dynamics of TCP over ATM in an environment with non-negligable cell losses. [Romanow 1995].

- The question:

Do we mostly care about the behavior of current implementations of TCP, including the limitations imposed by the coarse clock granularities? Or do we mostly care about the fundamental properties of the underlying algorithms?

- Answer:

It depends on the goal of that particular research.

Modeling a changing world: Reno TCP's unnecessary retransmit timeouts

- The issue:

Reno TCP has a problem with unnecessary retransmit timeouts when multiple packets are dropped from a window of data.
SACK TCP and NewReno TCP do not have this problem. [Fall 1996].

- The question:

Are we in danger of designing future networks (e.g., rate control algorithms for ABR service in ATM networks) tuned to the performance details of 1990 implementations of Reno TCP?

Modeling a changing world: traffic modeling

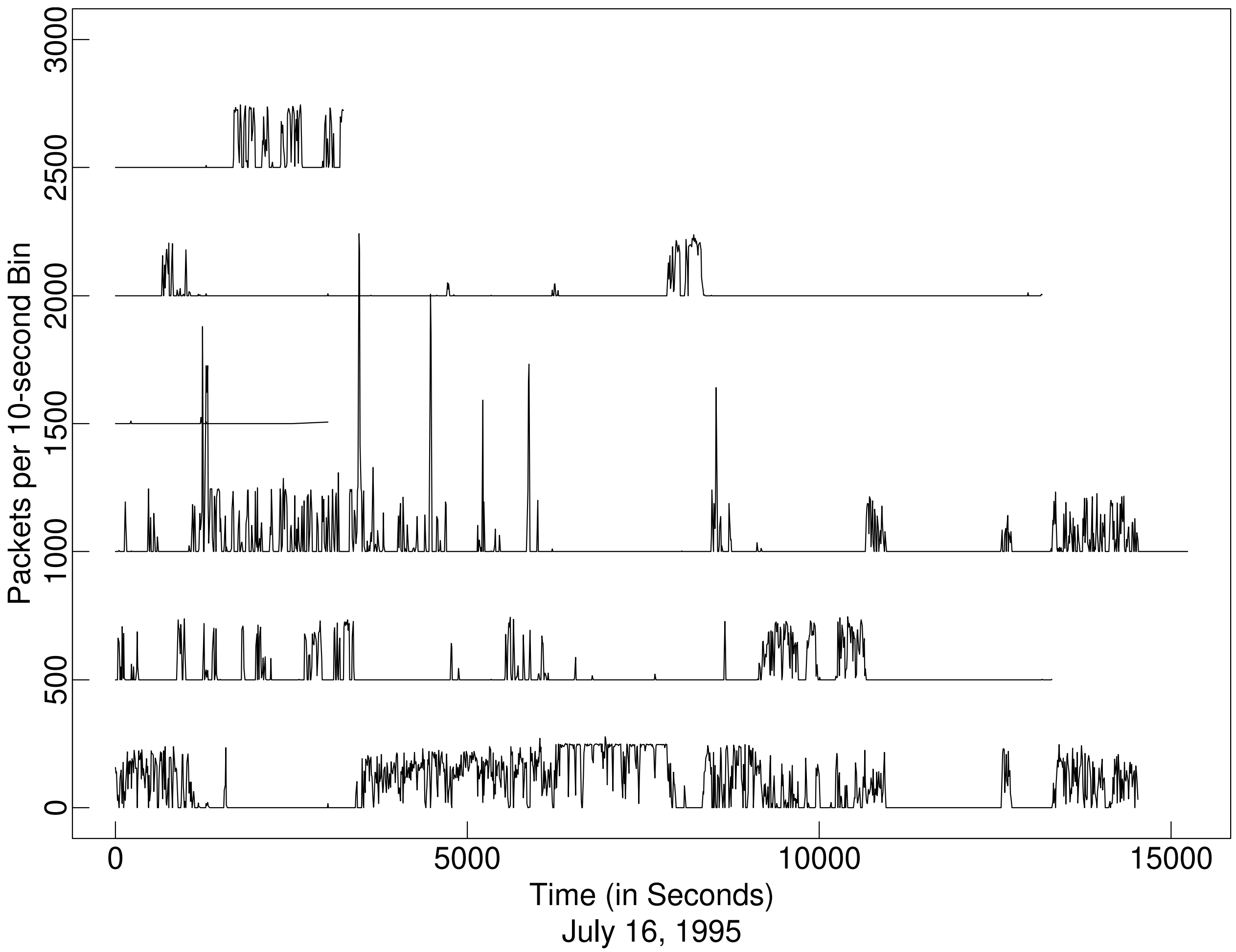
- The issue:

Measurement-based admissions control algorithms for realtime traffic with QoS requirements.

- The question of traffic mix:

Is the realtime traffic all MPEG video?
All two-way audio?
Or a mix that includes reliable multicast, n-way audio, multi-player games, distributed simulations, etc.

Traffic modeling: small audioconferences have different traffic patterns than two-way audio:



Six members, each with a peak rate of 64 Kbps. [Floyd 1996].

Measurement-based admissions control: more questions

- The mix between realtime and non-realtime traffic:

Is the realtime traffic allocated 100% or only 10% of the link bandwidth?

- Network questions:

What is the context: IETF's guaranteed service, controlled-load service, or predictive service (which is no longer standards-track)?
Another proposed Internet service? ATM?

- Application-level questions:

Does the realtime traffic have adaptive playback times, coding to maximize loss tolerance, and adaptive congestion control, or not?

Measurement-based admissions control: better tools can help.

• Libraries and models of traffic generators for a wide range of traffic types.

• Higher-level procedures for generating a rich traffic mix in a simulation.

• A simulator that includes new models for realtime services, new proposals for admissions control algorithms, and new models for adaptive congestion control for realtime traffic, so that researchers can explore how these various proposals interact with each other (and with other proposals for new multicast routing algorithms, new queue management algorithms, etc.)

Modeling a changing world: general questions.

- Topologies:

Single backbone? Multiple ISPs? Congested transoceanic links?

- Underlying technologies:

Multiaccess links? Wireless? Satellite? ATM clouds?

- Router functionality:

Scheduling mechanisms? Queue management? Multicast routing algorithms?

- Pricing structures?

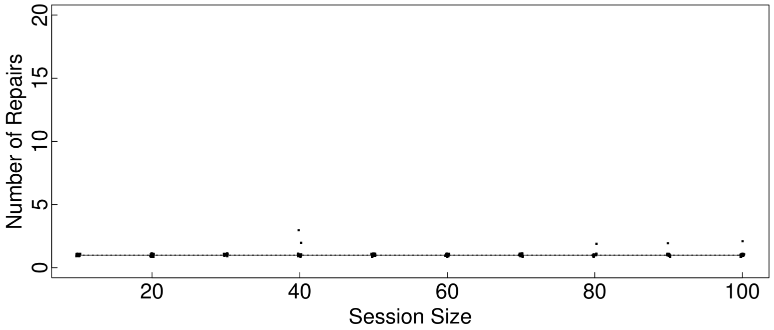
Issues of topology and of scale: Scalable Reliable Multicast.

• Session members in SRM use randomized algorithms for multicasting "requests" for missing data, and for multicasting "repairs" if they have the requested data. [Floyd 1995].

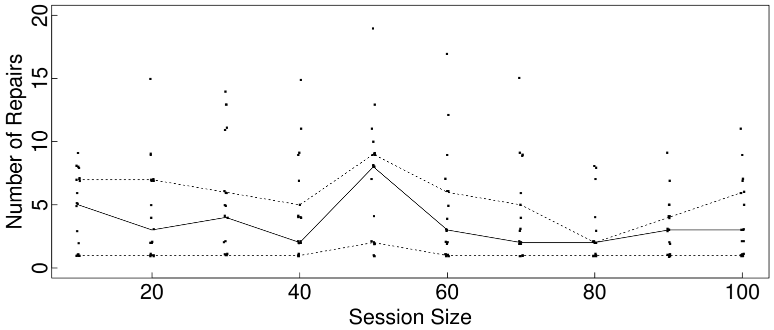
• But the results are very different for topologies of random labeled trees that only include session members, and for sessions embedded in topologies that are bounded-degree trees of degree four, with 1000 nodes.

SRM: results for the request/repair algorithms with fixed parameters.

- Random labeled graphs, all nodes are session members:

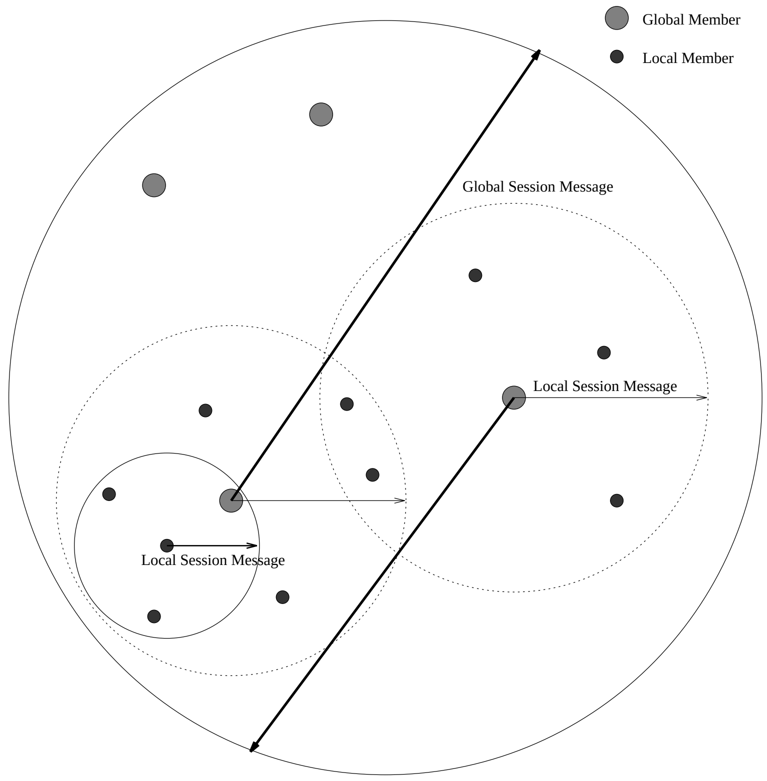


- Bounded-degree trees of degree four, 1000 nodes:

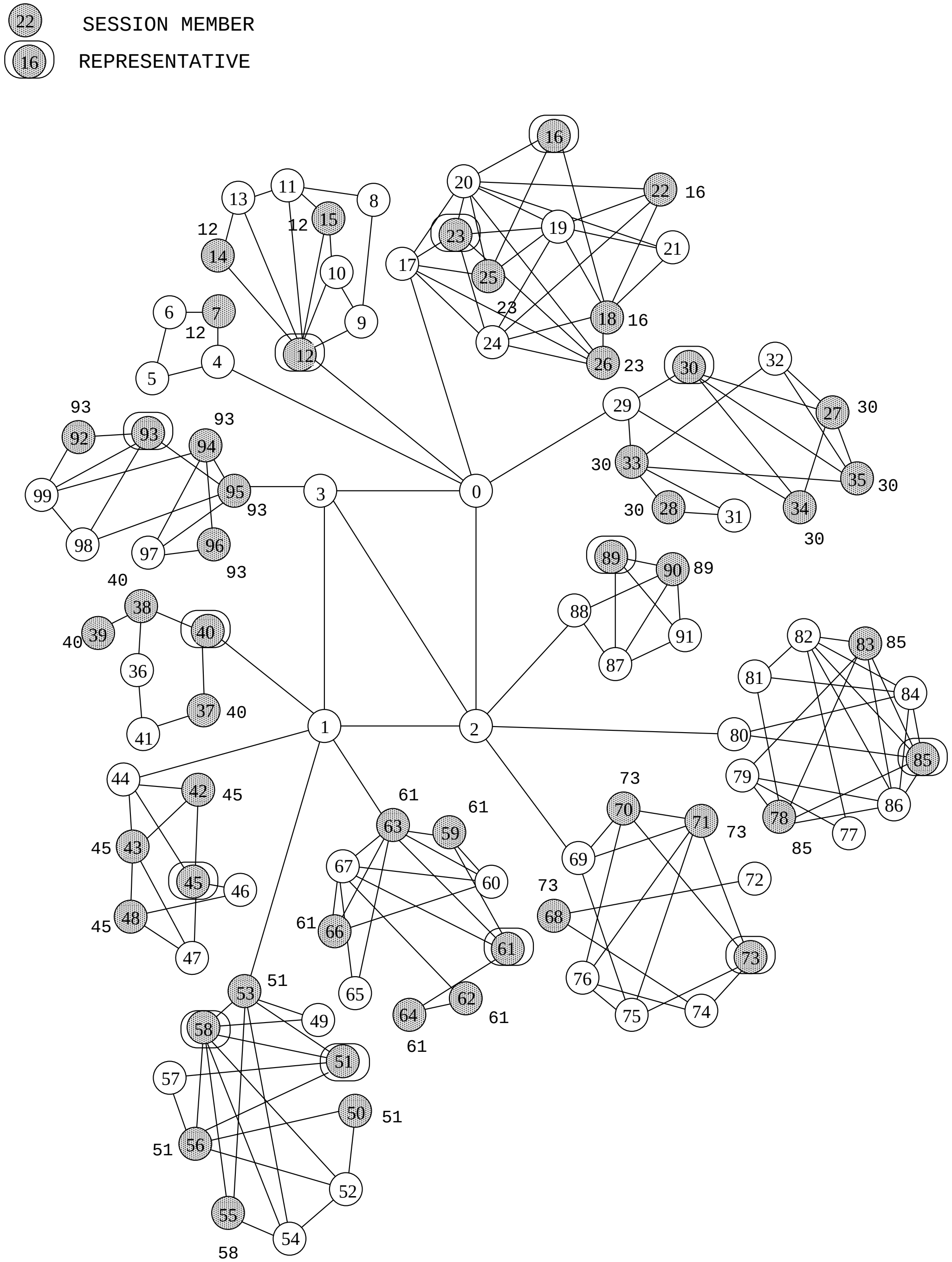


(The request/repair algorithms with adaptive parameters perform much better in this environment than the algorithms with fixed parameters.)

Issues of topology and of scale: Scalable Session Messages in SRM.



• Instead of all members sending periodic session messages with global scope, only “global representatives” send global session messages, and all members send local session messages. [Sharma 1997].



Scalable Session Messages:

- Issues of scale.
- Topology generation.
- Visualization tools?

Issues of topology and of scale: Reliable Multicast.

- Analysis of a complex scenario is generally intractable.

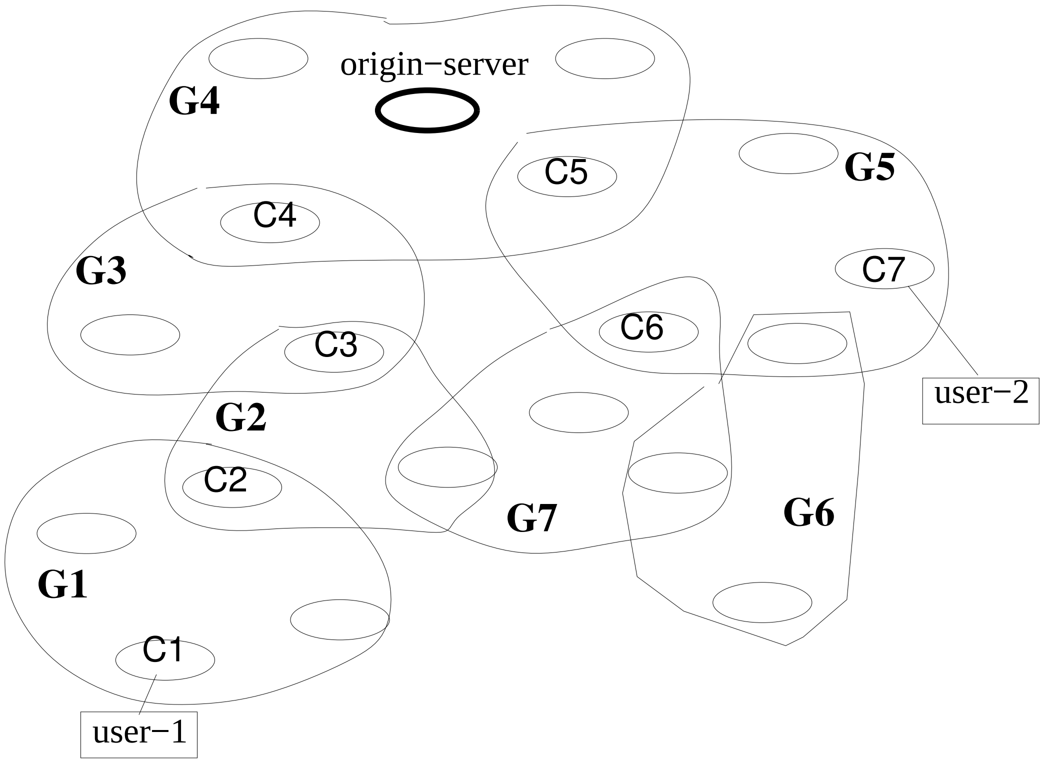
- Simulations need to include moderate-to-large multicast groups in a considerably-larger topology.

- Simulation results can be very sensitive to the underlying topology, so topology generation is an increasingly-important issue.

- We need visualization tools for large topologies that help us to develop intuition and understand the dynamics of the simulations.

Issues of topology and of scale: Adaptive Web Caching.

- A global data dissemination architecture: Web caches organize themselves into overlapping multicast groups. Web caches exchange information among themselves, and determine the most appropriate forwarding path for each request. [Zhang 1997].



Issues of topology and of scale: Adaptive Web Caching.

• Topology:
Probably the critical element of the underlying topology is that it consist of continents separated by congested high-delay inter-continental links. (In addition, the web servers are not evenly distributed among the continents.)

• Scale:
One question for a proposed mechanism should be at what scale does that mechanism begin to break down.
But how does one investigate in simulations a proposal for a global data dissemination architecture?

References (all available from http://www-nrg.ee.lbl.gov/floyd/):

[Fall 1996]: Fall, K., and Floyd, S., Simulation-based Comparisons of Tahoe, Reno, and SACK TCP, Computer Communication Review, V. 26 N. 3, July 1996.

[Floyd 1996]: Floyd, S., Comments on Measurement-based Admissions Control for Controlled-Load Services, July 1996.

[Floyd 1995]: Floyd, S., Jacobson, V., Liu, C., McCanne, S., and Zhang, L., A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing. SIGCOMM '95. To appear in Transactions on Networking.

[Romanow 1995]: Romanow, A., and Floyd, S., Dynamics of TCP Traffic over ATM Networks. IEEE JSAC, V. 13 N. 4, May 1995.

[Sharma 1997]: Sharma, P., Estrin, D., Floyd, S., and Zhang, L., Scalable Session Messages in SRM, August 1997, in submission.

References (continued):

[Zhang 1997]: Zhang, L,. Floyd, S., and Jacobson, V., Adaptive Web Caching, February 1997 proposal.

Wish list:

- Vint/ns and similar projects.
- Topology generators based on continents.
- Visualization tools.