Techniques for Simulation of Realistic Infrastructure Wireless Network Traffic

Caleb Phillips[1], Suresh Singh[2],
Douglas Sicker[1], and Dirk Grunwald[1]

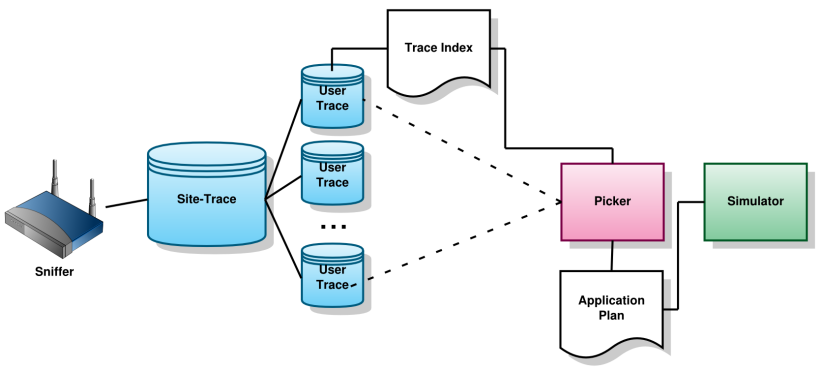
[1]Computer Science Department
University of Colorado, Boulder

[2]Computer Science Department
Portland State University

Problem and Approach

- **Problem:** In many cases, realistic simulation requires realistic background and/or foreground traffic
- **Approach:**
  - *Trace Replay:* Easiest way to generate realistic traffic is to directly use real traffic
  - *Trace Classification:* Need a way to intelligently choose from large corpus of available data
- **Validation:** How are application-layer results affected in a well-known application?
- **Scope:** Infrastructure Wireless Networks

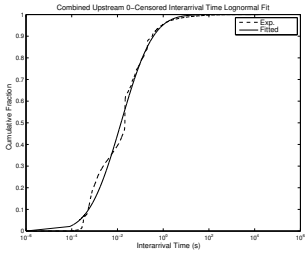
Framework

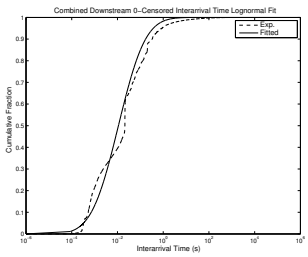

Trace Index
User Trace
User Trace
User Trace
Site-Trace
Sniffer
Picker
Simulator
Application Plan

Classifiers

A starting point for qualitative statistical classification...

- **Allan Deviation** as a measure of *burstiness*
- **IAT Lognormal Fits** as a measure of *traffic shape*
- **Stochastic Similarity** Clustering as a measure of *user similarity*





Case-Study/Validation - Bounded Slowdown

- **Bounded Slowdown**, proposed by Krashinsky et al.
- Backoff-based 802.11 powersaving scheme optimized for request/response http traffic.
- Validated using model-generated HTTP traffic (Mah et al.)
- **Question:** Would application-layer results differ if more diverse (and realistic) traffic were used in validation?

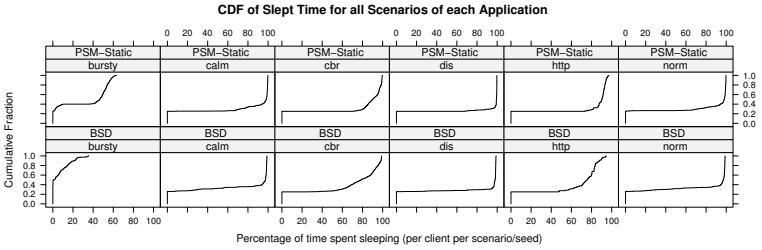
Data

294 unique, publicly available user-traces from these trace-sets:

- *PDX/Vwave2006:* Traces collected (by us) at public hotspots around Portland, Oregon
- *UW/SigComm2004:* Traces collected by UW researchers at SigComm 2004
- *Microsoft/OSDI2006:* Traces collected by Microsoft researchers at OSDI 2006

All traces are available at crawdad.org. Current software comes default with with 547 traces.

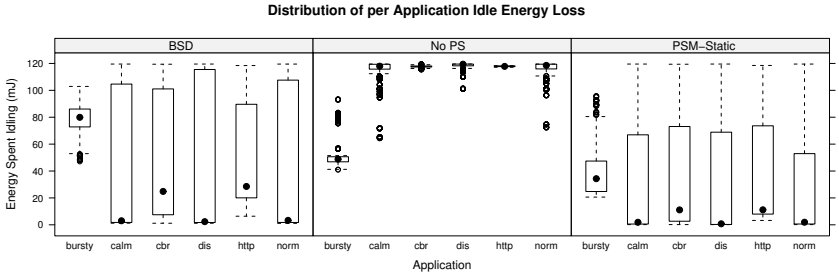
Simulation Results



ID	Type	Description
bursty	Trace Replay	Upper 10-percentile for Allan variance
calm	Trace Replay	Lower 10-percentile for Allan variance
dis	Trace Replay	Uniform across similarity clusters
norm	Trace Replay	Uniform random selection
cbr	Contrived	Randomly chosen between 1 and 40 packets/sec
http	Single App	HTTP traffic with a maximum think time of 1s

Table: Traffic sources used in analysis of Bounded Slowdown

Another Metric



Key Observation: Application-specific and contrived sources do not capture the dynamics of real traffic. This is especially apparent in pathological scenarios such as bursty users.

Analysis

- **Question:** Why is *http* so different, especially for bursty users?
- Generated HTTP traffic: IAT fits Generalized Pareto, mean Allan deviation of 5.62 for downstream and 3.32 for upstream (versus 1.16/0.90 in real traces), stochastic similarity minimum is 0.99.
- Just one application but results are motivating, particularly in terms of ease and flexibility of the framework...

Future Work

- Other (better?) trace classification metrics
- More validation, additional applications
- Expand to other network types
- Maintain and expand implementation: more data, more simulators

Questions?

Contact: caleb.phillips@colorado.edu
Download: http://systems.cs.colorado.edu

Limitations and Solutions

- Implications of replaying "shaped" traces (Floyd et al.)
- Introduction of SWING or similar (Vishwanath et al.) into framework