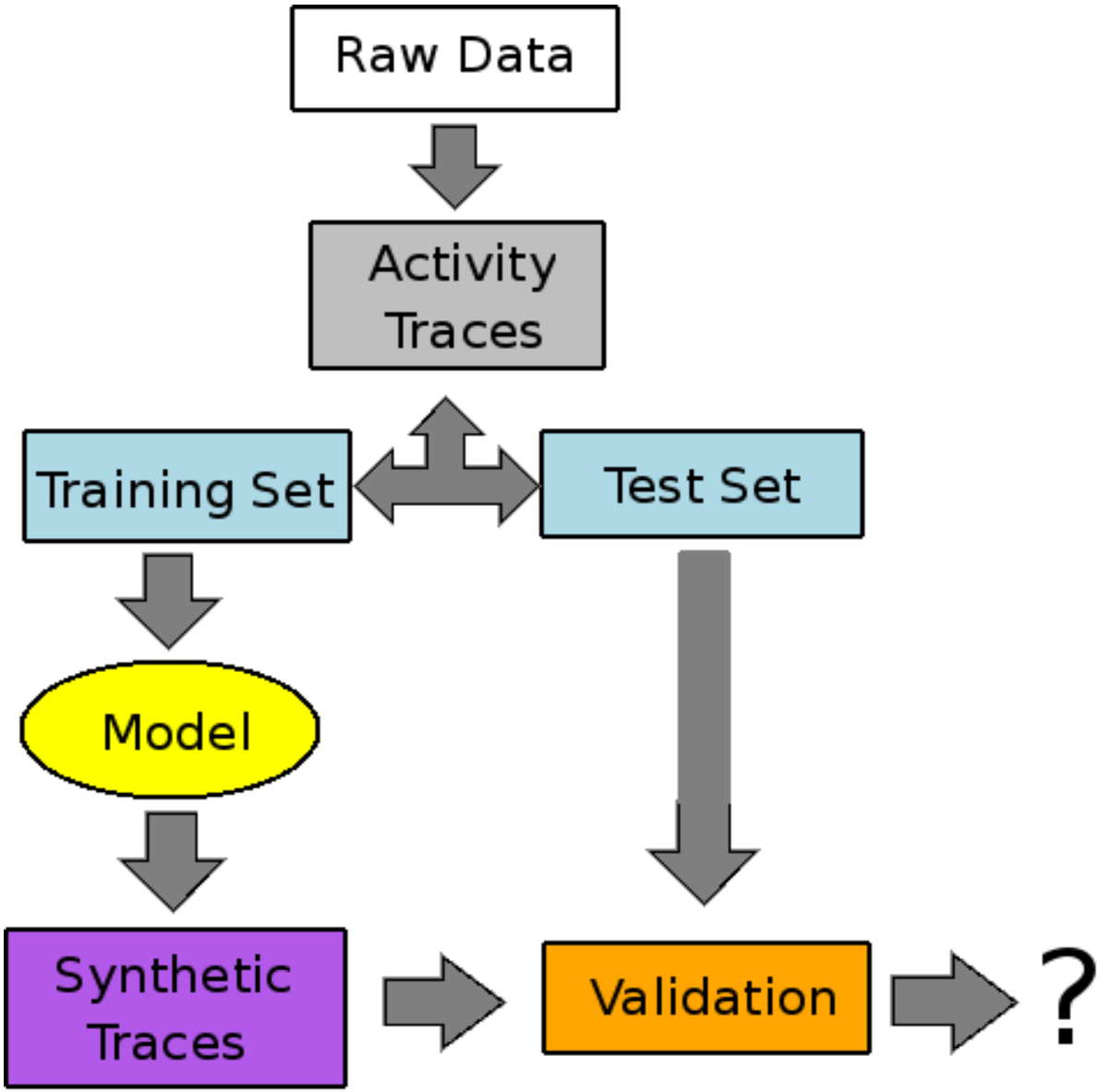
An Empirical Activity Model for WLAN Users

Caleb Phillips[1] and Suresh Singh[2]
[1]Computer Science Department, University of Colorado, Boulder
[2]Computer Science Department, Portland State University
caleb.phillips@colorado.edu, singh@cs.pdx.edu

Problem

- Develop a simple, measurement-driven, and protocol independent model for WLAN user activity.
- Rigorously validate the effectiveness of this model.
- Such a model would be useful for applications such as power-saving and load-balancing.

Outline


Raw Data
Activity Traces
Training Set
Test Set
Model
Synthetic Traces
Validation
?

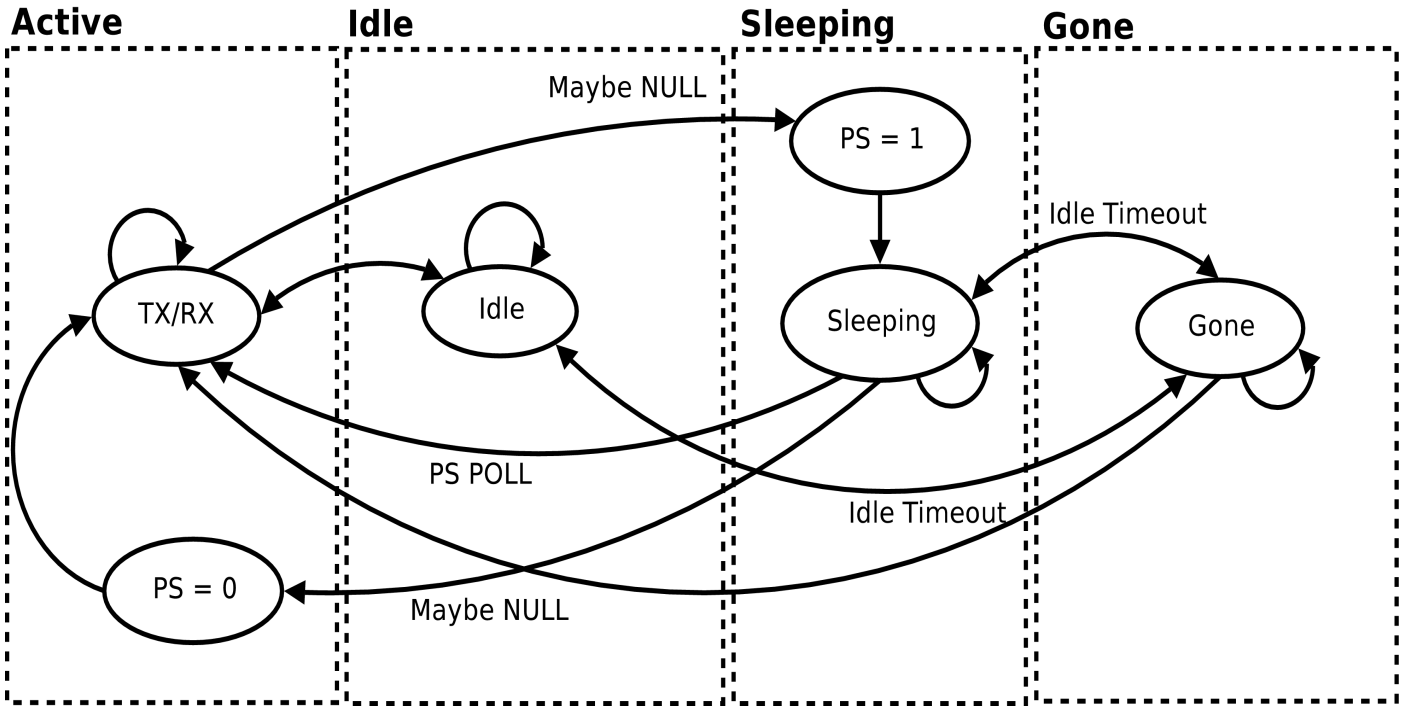
Data

- 5 locations, 131 distinct users:
  - 4 collected at public hotspots around Portland, Oregon [Phillips'07]
  - 1 collected by UW researchers at SigComm'04 [Rodrig'05]

- All traces are available on CRAWDAD.

CRAWDAD

Data => Activity Traces



- Map 802.11 actions to four discrete states: Active, Idle, Sleeping, or Gone.
- Packets are grouped into one second buckets.

Deriving and Using a Model

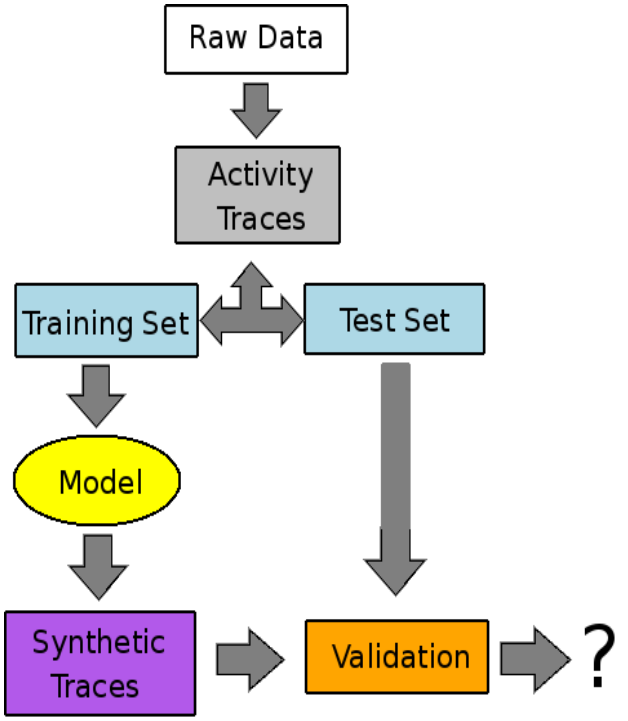
Transition Probability Matrix
Sleepless Category
Tx

	active	idle	sleeping	gone
active	0.85	0.15	0	0
idle	0.05	0.95	0	0
sleeping	0	0	0	0
gone	0	0	0	1

- Users are categorized as sleepy or sleepless.
- Determine per-category transition probabilities
- Trace Length from Pareto-distributed fit to Dartmouth data.
- Synthetic traces created via Markov simulation

Validation

- Cross validation
- No straw-man...
- Synthetic user traces are compared to test set using a stochastic similarity metric [Nechyba'98]



$\sigma(\blacktriangle, \bullet) = ?$

Stochastic Similarity

- EM iteratively maximizes the probability of the i[th] model (λ_i) given the i[th] observations (O_i):

$$P(\lambda_i | O_i)$$

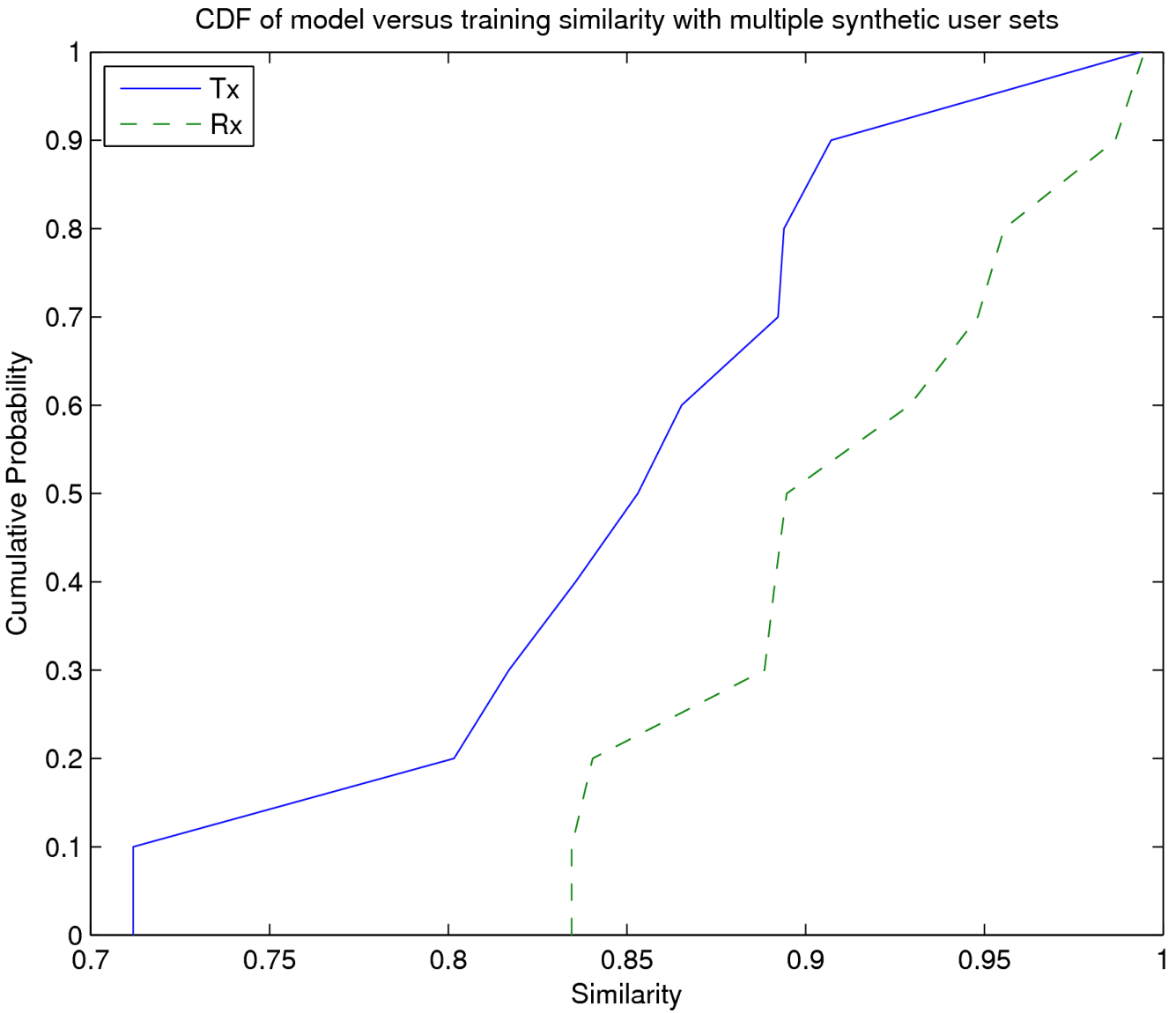
- The likelihood of a given observation given some model scaled by the observation length (T_i):

$$P_{ij} = P(O_i | \lambda_j)^{1/T_i}$$

- Similarity between observations:

$$\sigma(O_i, O_j) = \sqrt{\frac{P_{ji} P_{ij}}{P_{ii} P_{jj}}}$$

Validation Results (a Selection)



Conclusions

- A simple, static, four-state Markov model is sufficient to model user activity at the granularity we have chosen to approach.
- User activity exhibits alternating periods of activity and idle time, often several seconds long.
- Stochastic similarity is a useful tool for model validation.

More Conclusions

- Categorization of users can improve model fidelity.
- Inter-trace similarity comparison shows that seemingly different venues yield similar traces.

Ongoing Work

- Application of the model developed here to dynamic power-saving applications for WLAN devices [appearing at MobileHCI'08]
- In future work, we would like to further investigate the application of similarity classification to a wide variety of measured trace-sets.

Questions?

- Offline questions/comments?

  caleb.phillips@colorado.edu

- Our data is available on CRAWDAD:

  http://crawdad.cs.dartmouth.edu/meta.php?name=pdx/vwave

- Similarity implementation (Matlab) is available for download:

  http://www.cs.pdx.edu/~singh/software/similarity-0.1.tar.gz