

# Research Methods in computer science

Fall 2013

Lecture 9

Omprakash Gnawali  
September 26, 2013

# Agenda

Research Conference Updates

Sampling Bias

HW4 and HW5 feedback

HW6

# Sampling Bias

“In statistics, sampling bias is a bias in which a sample is collected in such a way that some members of the intended population are less likely to be included than others. It results in a biased sample, a non-random sample of a population (or non-human factors) in which all individuals, or instances, were not equally likely to have been selected. If this is not accounted for, results can be erroneously attributed to the phenomenon under study rather than to the method of sampling.” -- wikipedia

## Types of sampling bias

Self selection bias

Pre-screening

Exclusion

etc.

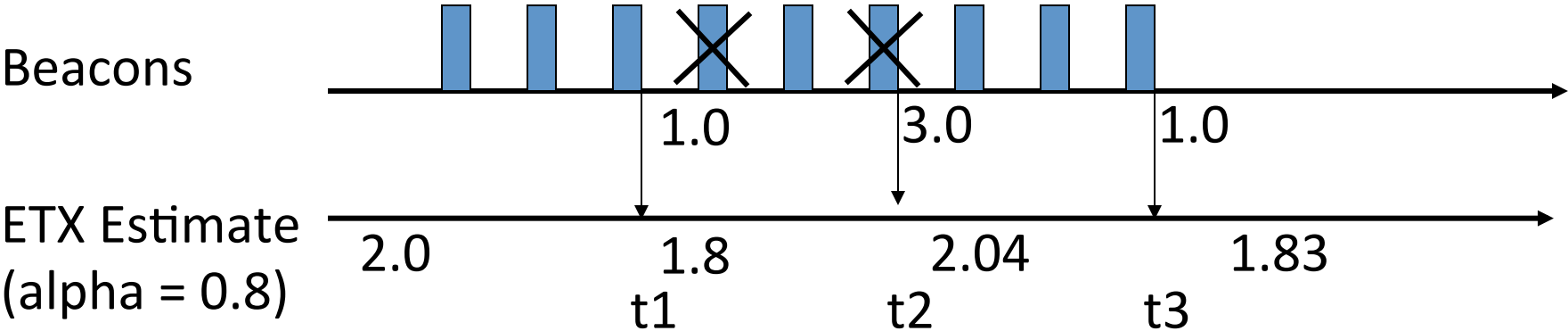
[from wikipedia]

Using signal strength for link quality estimation can introduce sampling bias.

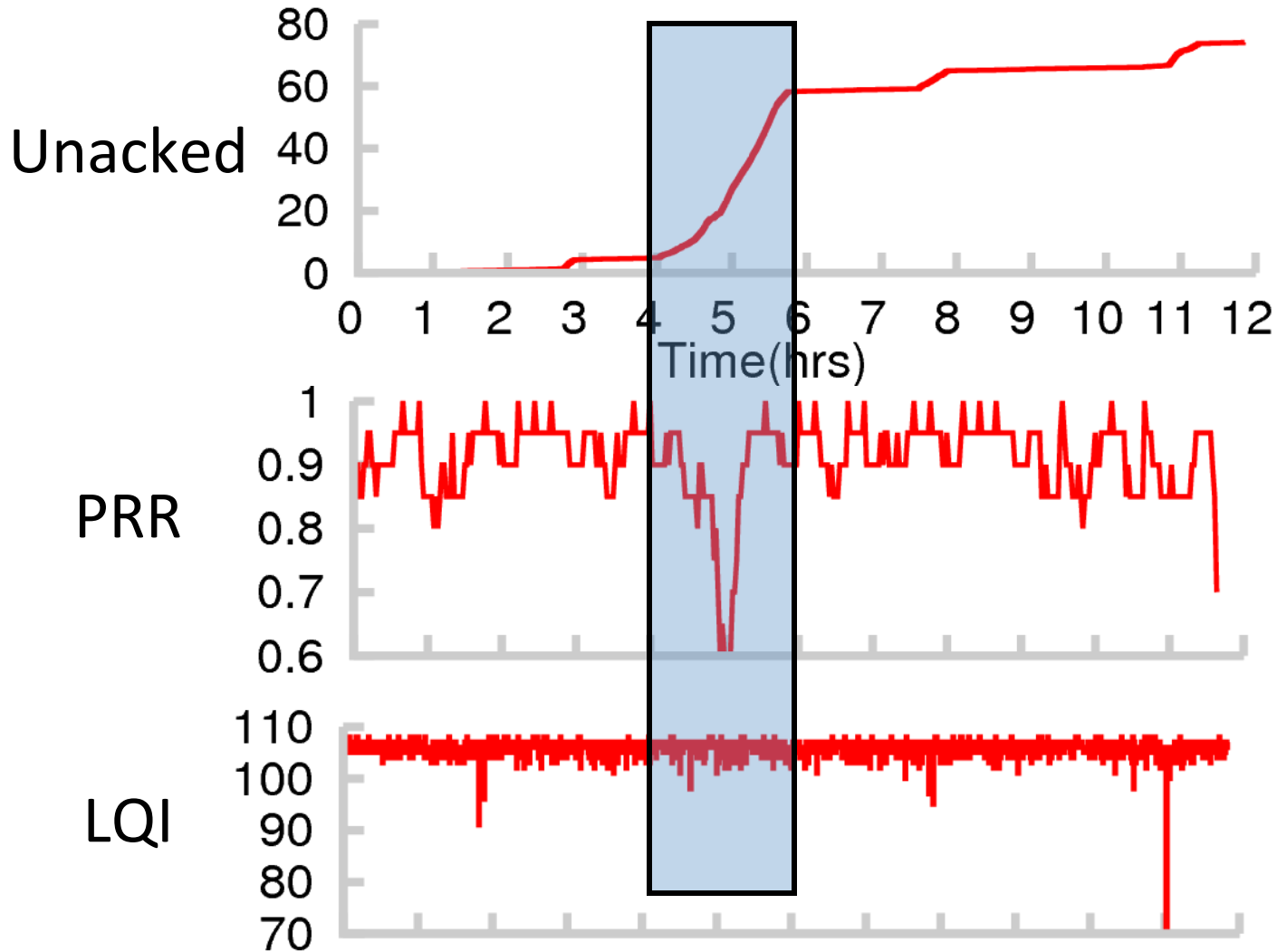
## Link quality estimation

Estimate how “good” a link is.  
Important for link selection.

# ETX Estimation Example



# Link Estimation using PHY info



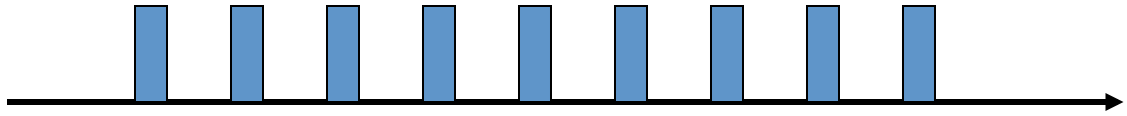


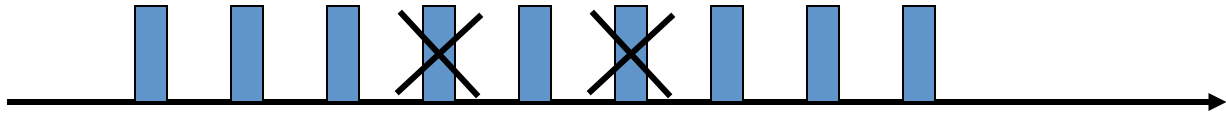
Quality of reception = Signal / Noise

Lot of wireless network research tries to understand performance as some function of SNR

“The Prism 2.5 chip-set provides per-frame measurements called RSSI (receive signal strength indication) and “silence value.” The RSSI reflects the total power observed by the radio hardware while receiving the frame, including signal, interference, and background noise. The silence value reflects the total power observed just before the start of the frame. We found that the accuracy of the RSSI and silence readings was within 4 dB by comparison with a spectrum analyzer. This paper reports signal-to-noise ratios derived from the RSSI and silence values.” – [Aguayo et al. 2004]

Quality of reception = Signal / Noise



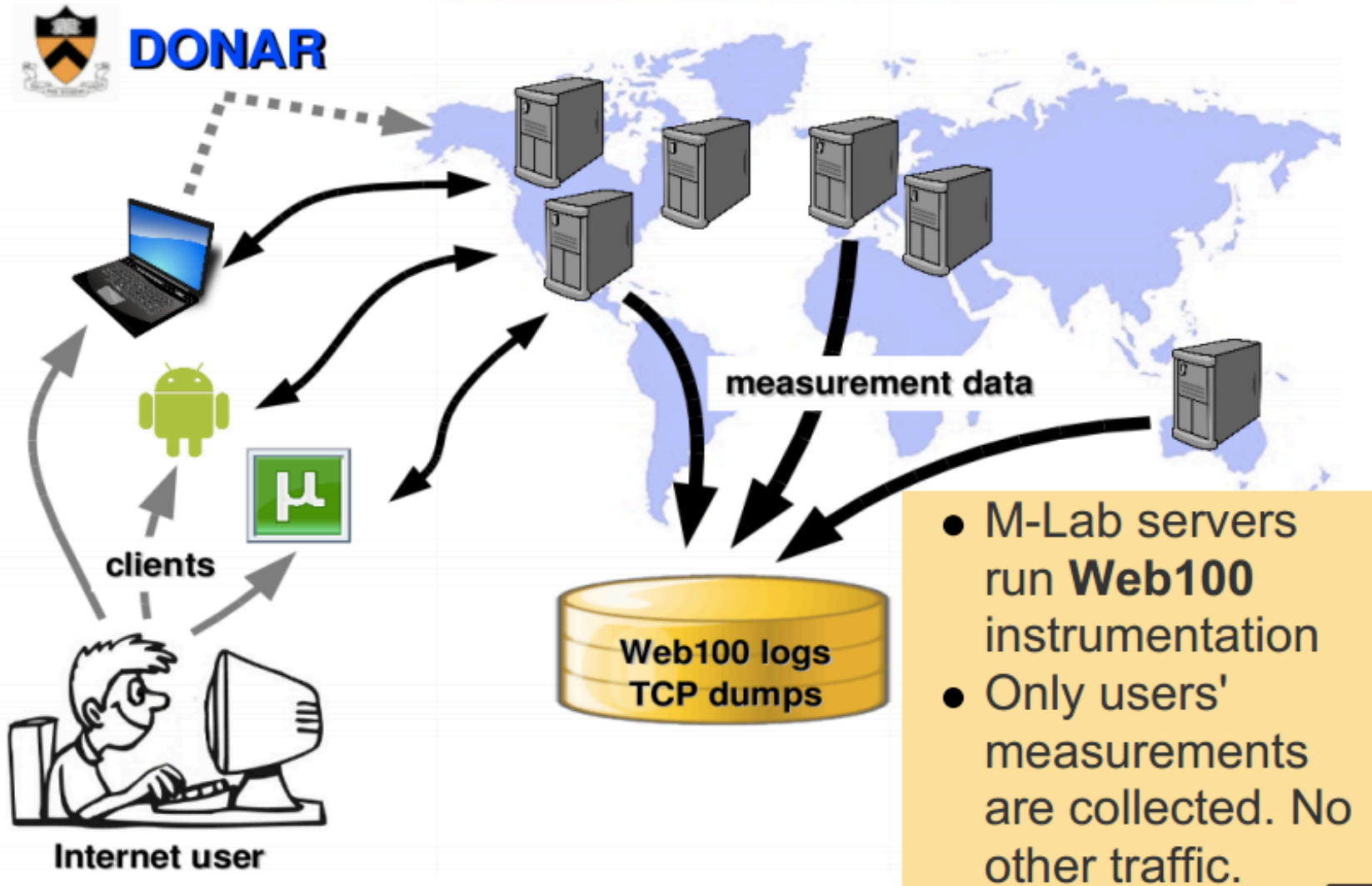


How to select participants  
for an HCI study?

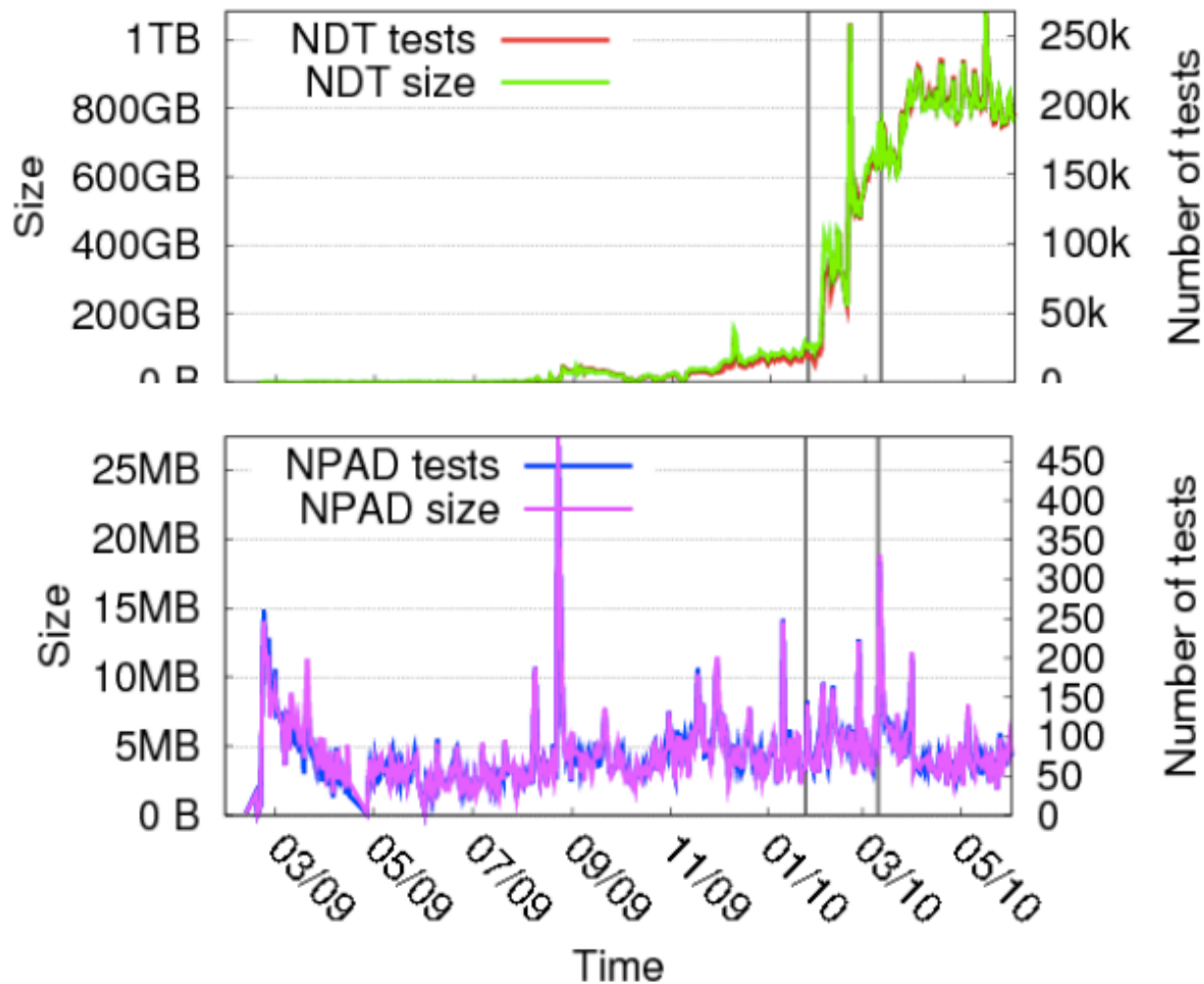
What is the possibility of  
sampling bias?

Example of bias due to a  
significant change in mixture of  
data source

# Measurement & Data collection



# How much data? How many tests?



Jan 25 2010  
uTorrent launch  
Mar 11 2010  
FCC launch

**NDT**

*Tot tests: 22M*

*Tot size: 93TB*

**NPAD**

*Tot tests: 34K*

*Tot size: 2GB*



# HW5

Write a response to the pamphlet on smart meters.

# HW6

Make a list of metrics from the ten research papers you selected

Define each metric in a few sentences.