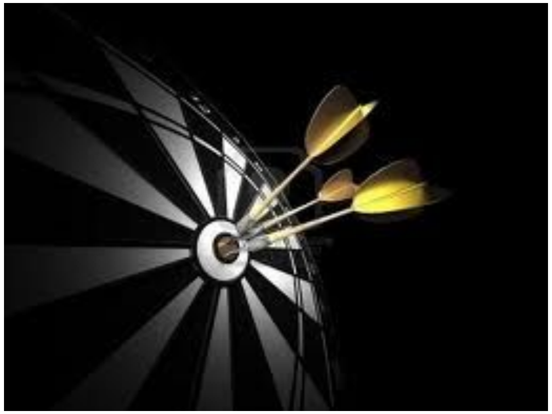


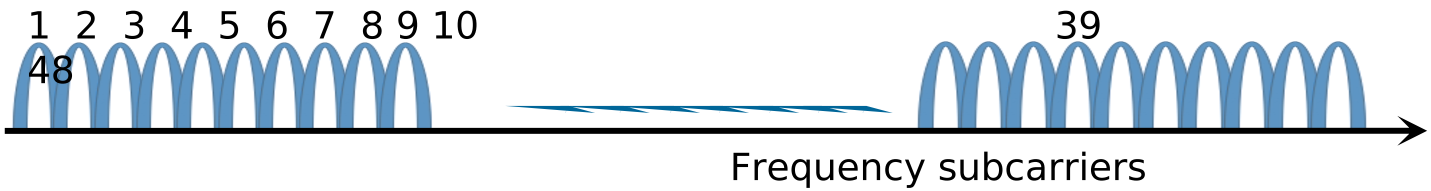
Precise Indoor

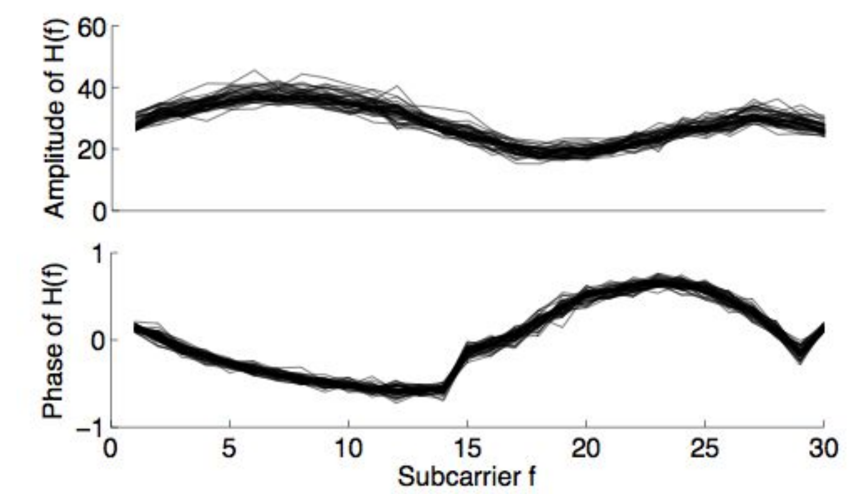
Spot localization using PHY Layer information

Mobisys' 12

Fingerprinting Wireless Channel

- 802.11 a/g/n implements OFDM
  - Wideband channel divided into subcarriers





  - Intel 5300 card exports frequency response per subcarrier

Is WiFi Channel Amenable to Localization?

- Two key hypotheses need to hold:

1. **Temporal**
    - **Channel responses at a given location may vary over time**
    - **However, variations must exhibit a pattern – a signature**

2. **Spatial**
    - **Channel responses at different locations need to be different**

Experimental methodology

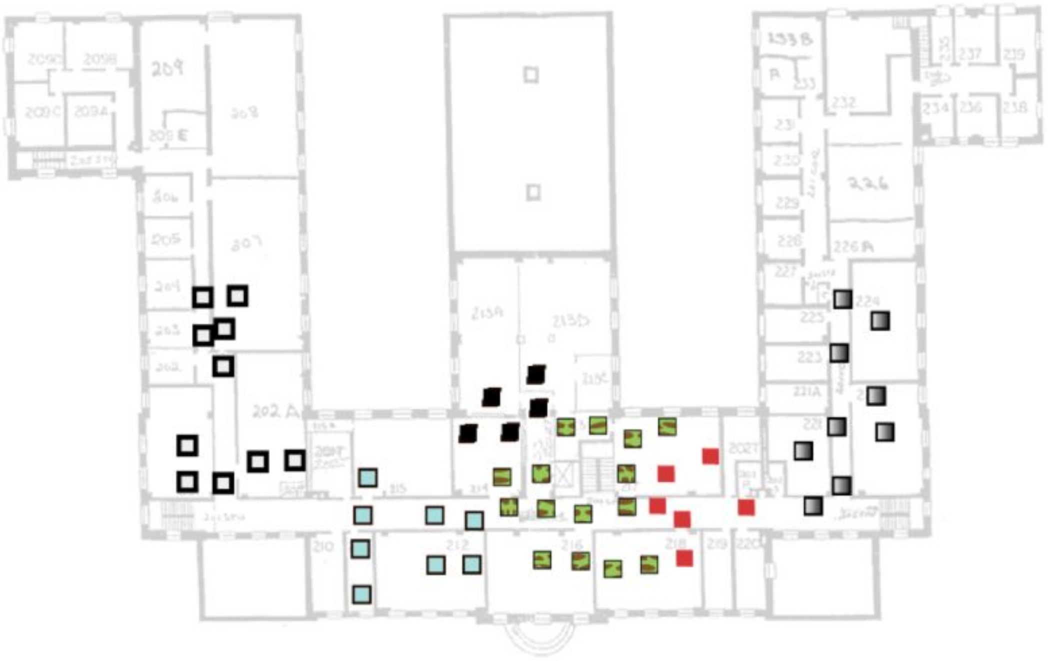
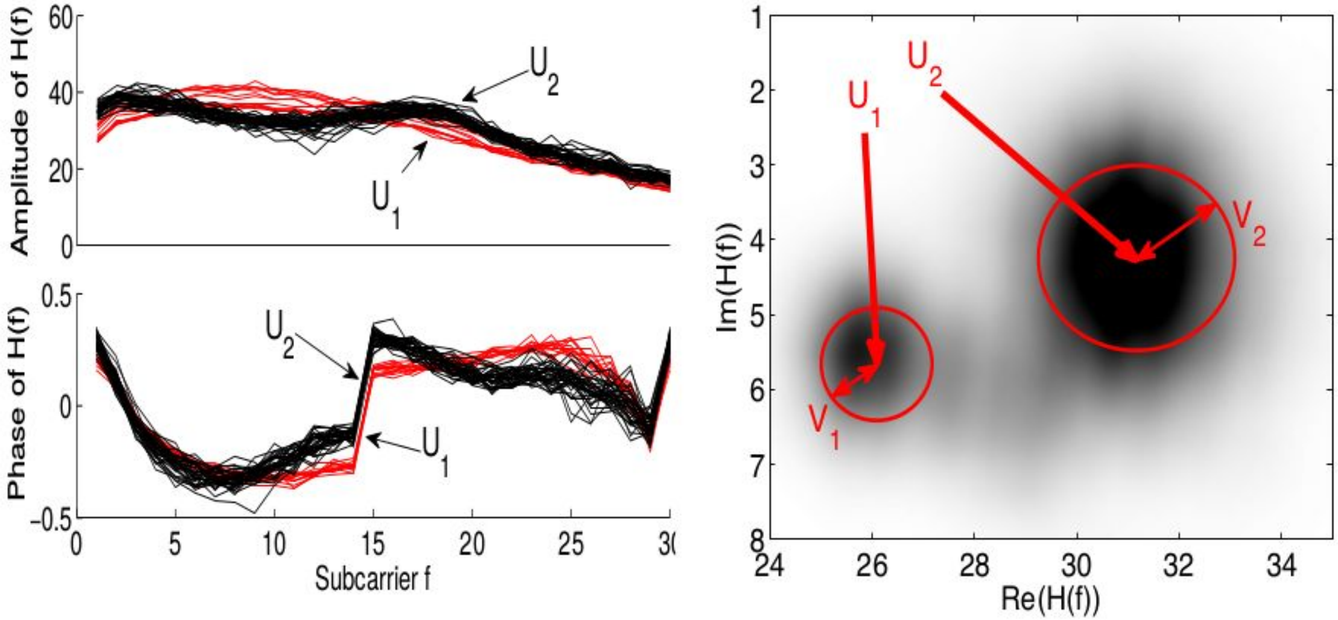


Figure 1: Engineering building floorplan. Different sets of spots shown in different colors – our initial measurements in this section uses only one set of 15 spots (shown in green).

Variation over Time

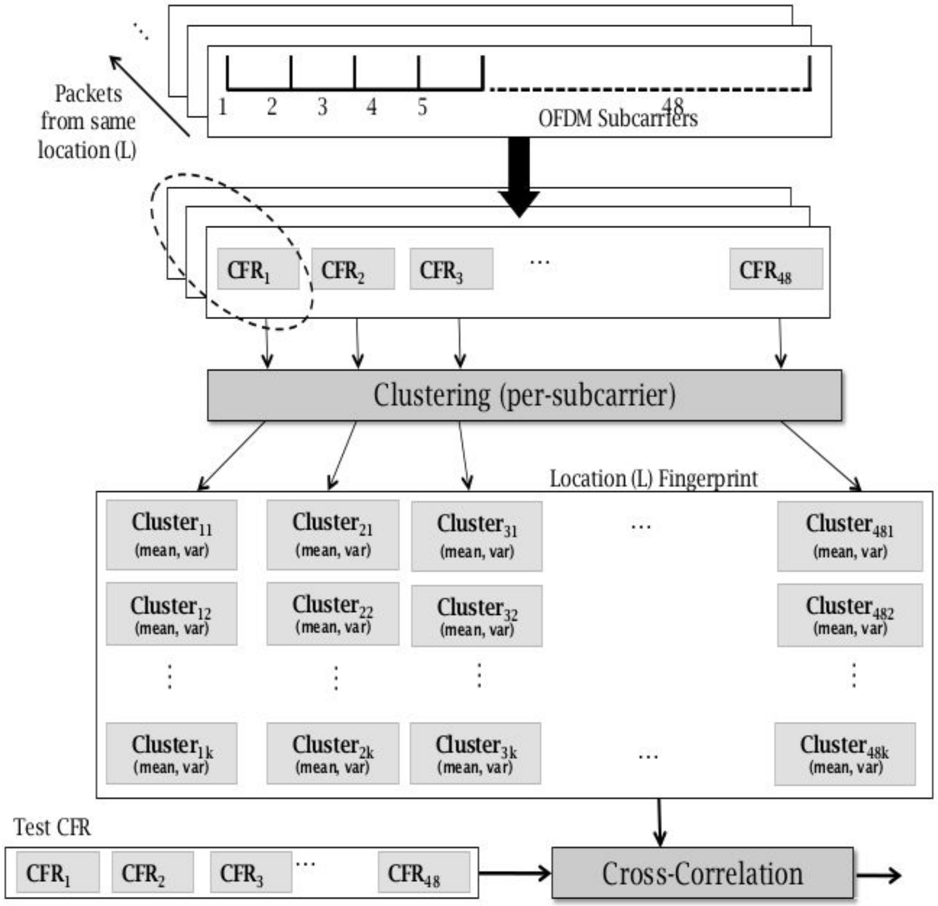
- Measured channel response at different times



Observe: Frequency responses often clustered at a location

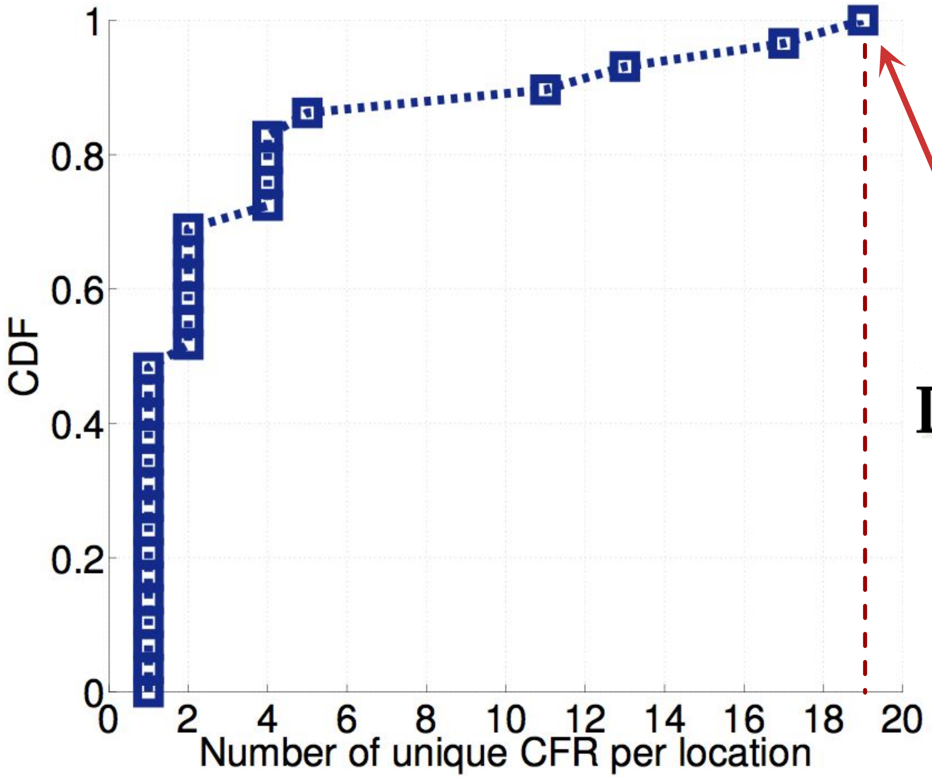
But not necessarily one cluster per location

Overview



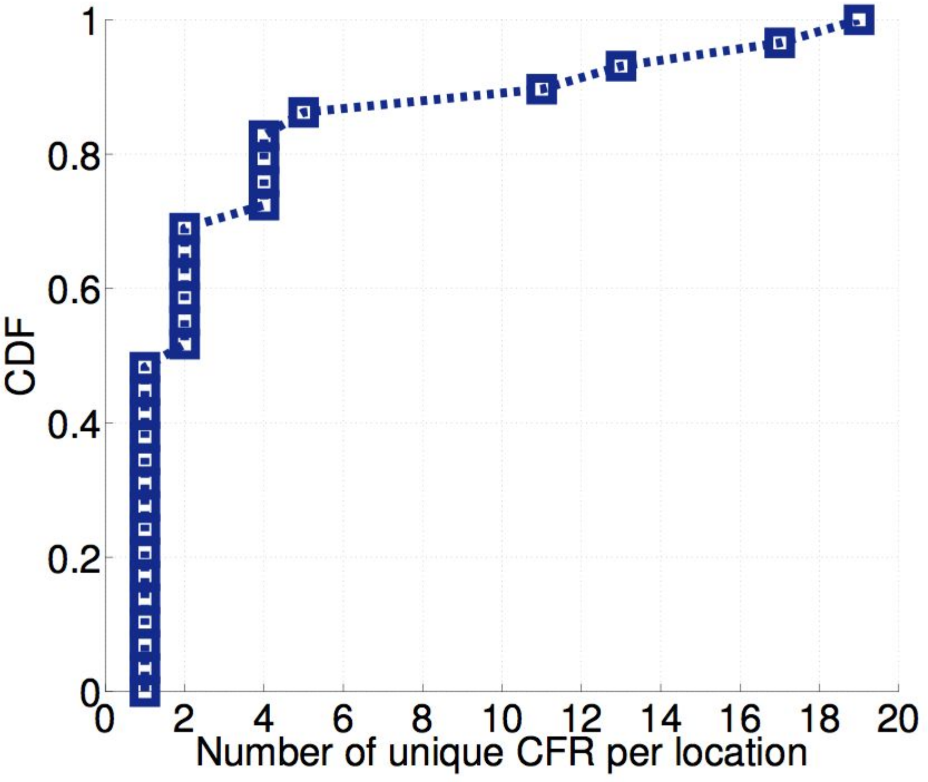
Packets
from same
location (L)
1 2 3 4 5 48
OFDM Subcarriers
CFR1 CFR2 CFR3 … CFR48
Clustering (per-subcarrier)
Location (L) Fingerprint
Cluster11 (mean, var)
Cluster21 (mean, var)
Cluster31 (mean, var)
Cluster481 (mean, var)
Cluster12 (mean, var)
Cluster22 (mean, var)
Cluster32 (mean, var)
Cluster482 (mean, var)
Cluster1k (mean, var)
Cluster2k (mean, var)
Cluster3k (mean, var)
Cluster48k (mean, var)
Test CFR
CFR1 CFR2 CFR3 … CFR48
Cross-Correlation

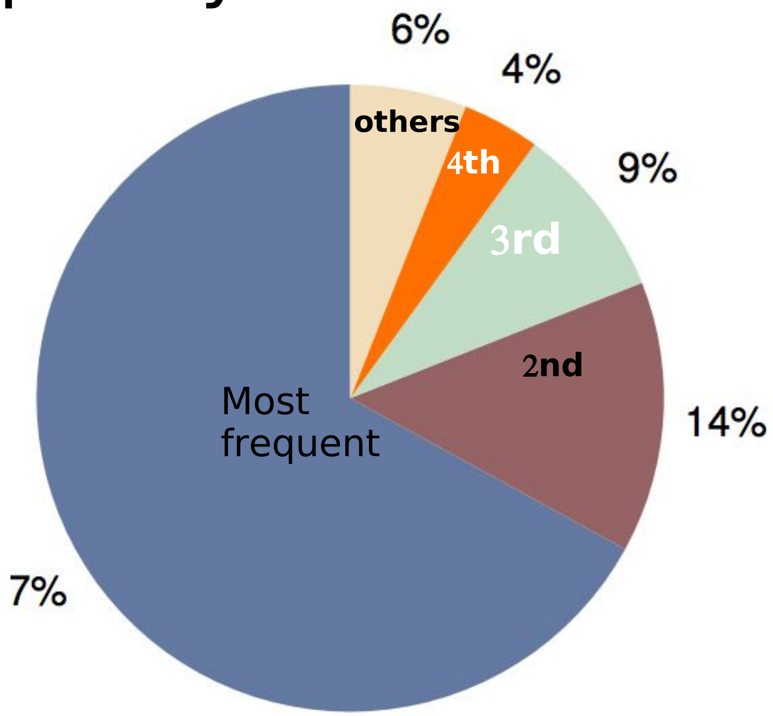
How Many Clusters per Location？



Do all 19 clusters occur with same frequency?

Cluster Occurrence Frequency





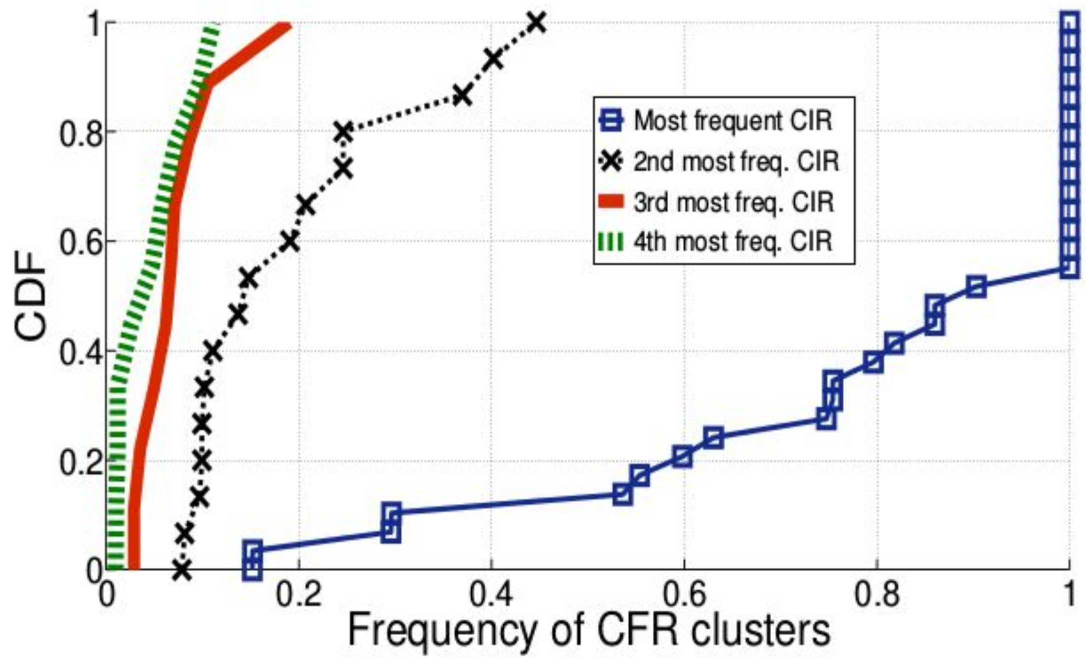


Figure 4: (a) CDF of the number of CFR clusters observed at 30 different client locations (i.e., 30 distinct links); (b) CDF of the probability of seeing the n-th most frequent CFR cluster.

Impact of environmental changes

Measure the cross-correlation of CFR observed with and without the change

$$c(\mathbf{a}, \mathbf{b}) = \frac{\sum_i a_i b_i}{\sqrt{\sum_i a_i^2}\sqrt{\sum_i b_i^2}}$$

Impact of environmental changes

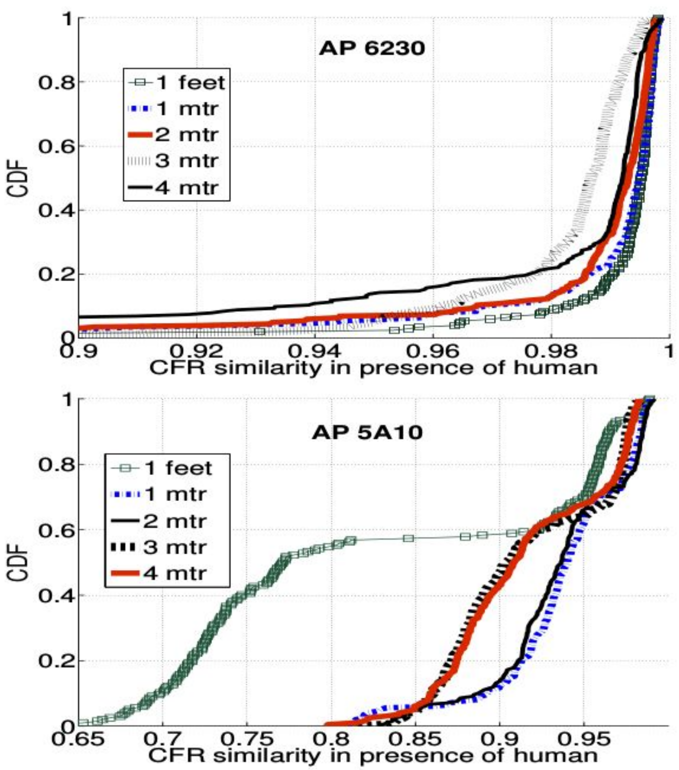


Figure 5: CFR cross correlation in presence of human beings at a location for 2 different APs at 2.4GHz

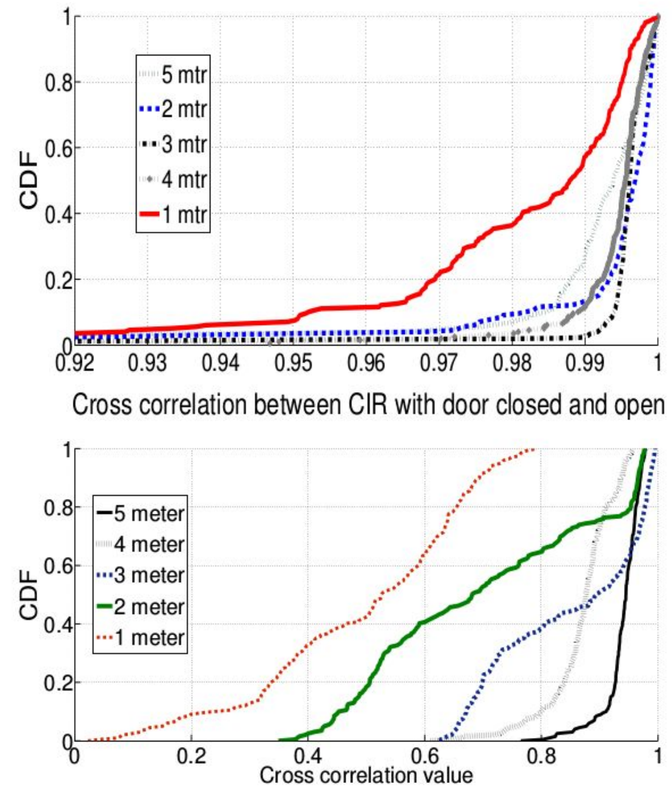
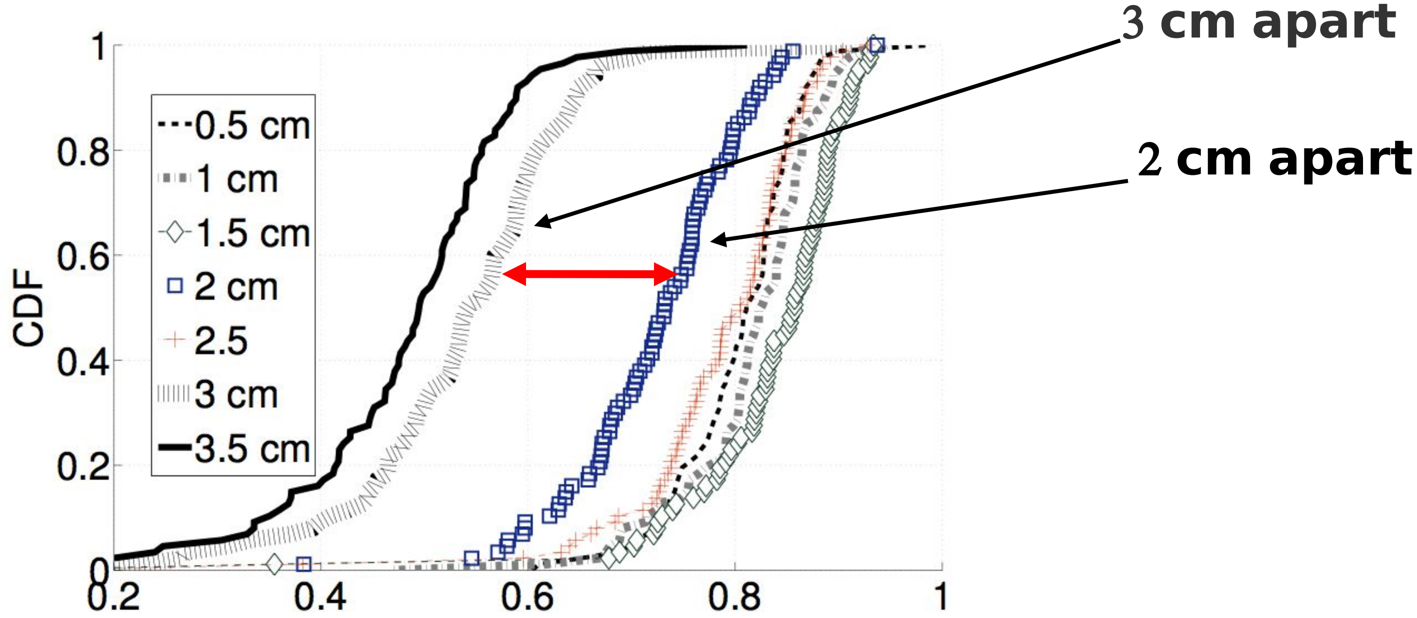


Figure 6: (a) CFR cross correlation for (a) door open vs. closed and (b) original metal shelf vs. moved; for various distances from the laptop.

What is the Size of a Location？

- Localization granularity depends on size
  - RSSI changes in orders of several meters（hence, unsuitable）



Cross correlation with signature at reference location

CFR structure varies from one location to another

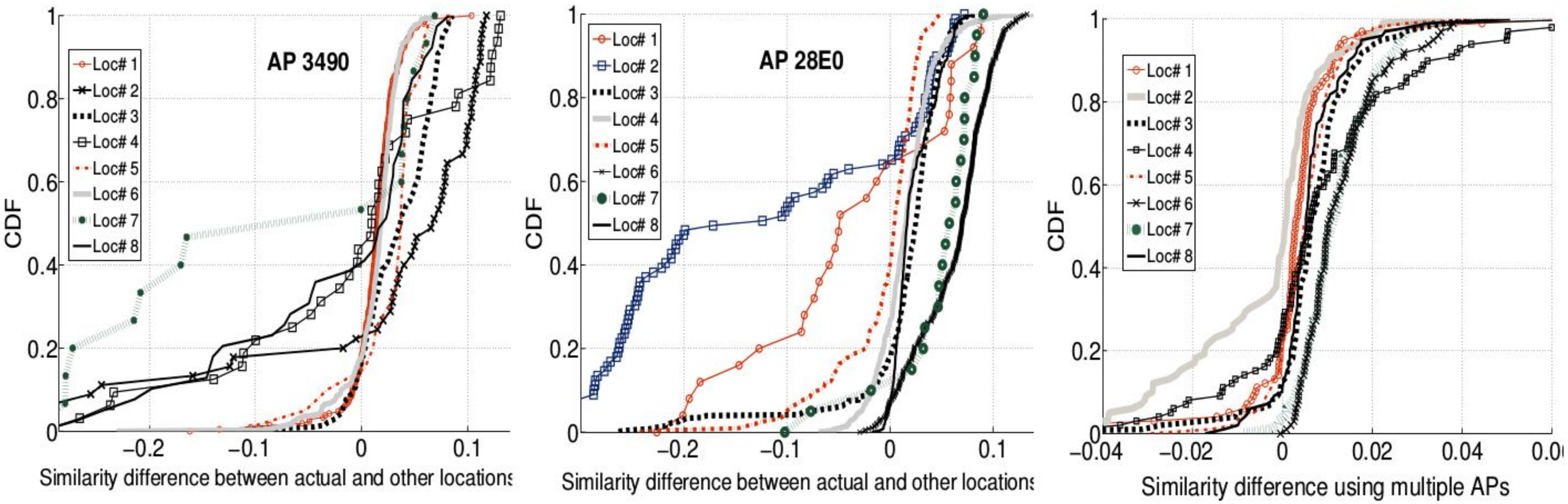


Figure 7: CDF of the difference in similarities $S_{own} - S_{others}$ observed at 8 locations, for two different access point: (a) AP 3490, (b) AP 28E0. (c) CDF of the maximum similarity difference ($S_{own} - S_{others}$) across all APs.

Pixels (locations) and spots

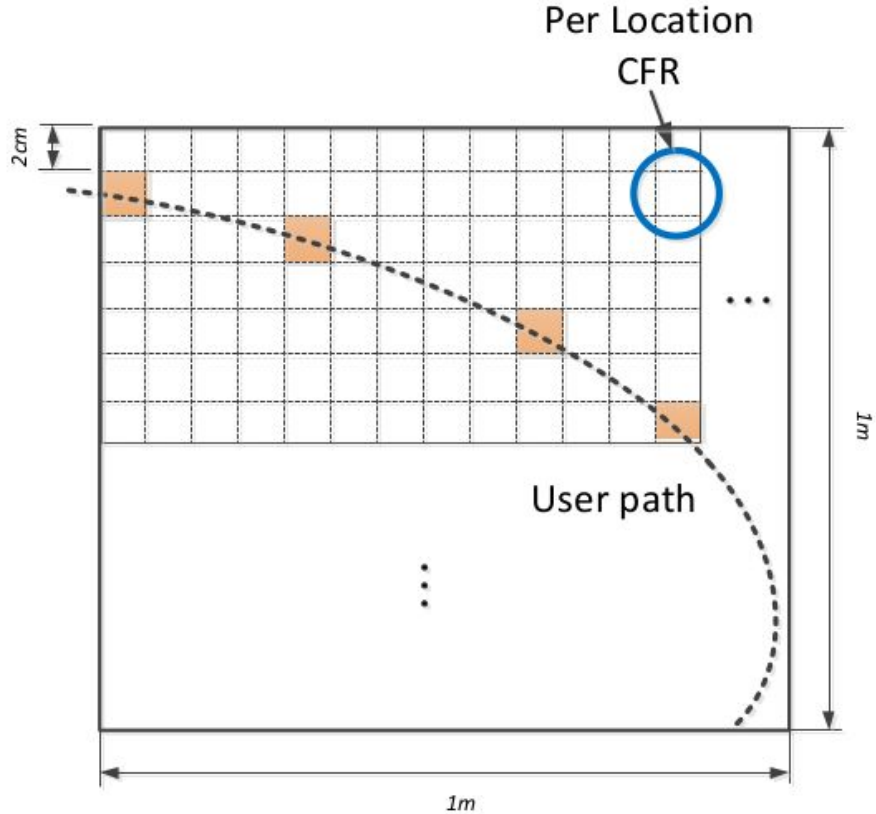


Figure 10: A device records multiple CFRs from a spot.

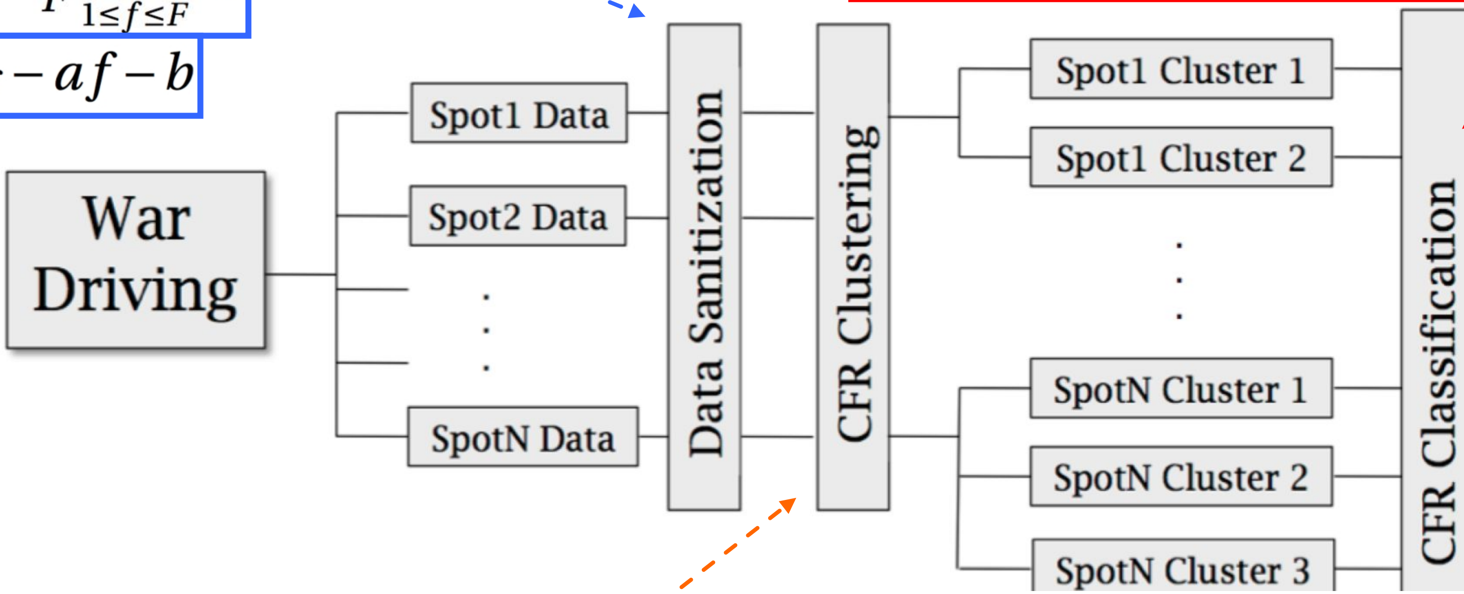
Architecture and Modeling

$$\hat{\phi}_f = \phi_f + 2\pi f_f \Delta t + \beta + Z_f$$

$$a = \frac{\hat{\phi}_F - \hat{\phi}_1}{2\pi F}, \qquad b = \frac{1}{F} \sum_{1 \le f \le F} \hat{\phi}_f$$

$$\hat{\phi}_f - af - b$$

$$d(\mathbf{P}, \mathbf{U}^i) = \sum_{f=1}^{F} \log(V_f^i) + \sum_{f=1}^{F} \left(\frac{\|P_f - U_f^i\|^2}{\left(V_f^i\right)^2} \right)$$



Parameters: (w_K, U_K, V_K)

Variational Inference (Infer.NET)

Test data

$$d(\mathbf{P}, S_i) = \min_{\mathbf{U}^i \in Z_i, AP(\mathbf{U}^i) = AP(\mathbf{P})} d(\mathbf{P}, \mathbf{U}^i)$$

Data sanitization

CFRs received at a location cannot be directly used for calibration.
Unknown phase and time lag can distort CFR.
We need to make sure that every the measurement includes same values of phase and time lag.

$$\boxed{\hat{\phi}_f = \phi_f + 2\pi f_f \Delta t + \beta + Z_f}$$

$$\boxed{\begin{aligned} a &= \frac{\hat{\phi}_F - \hat{\phi}_1}{2\pi F}, \\ b &= \frac{1}{F} \sum_{1 \le f \le F} \hat{\phi}_f \end{aligned}}$$

$$\boxed{\hat{\phi}_f - af - b}$$

Modeling channel response

- Model the noise as complex Gaussian noise.
- Model the channel response as a random vector with Gaussian mixture distribution.
- Channel response is assumed to be drawn from one of the representative CFR clusters chosen at random for each packet.
- Each CFR cluster is modeled as a complex Gaussian random vector with mean Ui and variance Vi.
- Probability that packet P belongs to CFR cluster with mean Ui

$$\mathrm{P}(\mathbf{P}|\mathbf{U}^i, \mathbf{V}^i) = \prod_{f=1}^{F} \frac{1}{2\pi \left(V_f^i\right)^2} \exp\left(-\frac{||P_f - U_f^i||^2}{2\left(V_f^i\right)^2}\right).$$

Applying logarithm and remove constants
to derive the loglikelihood distance metric.

$$\boxed{d(\mathbf{P},\mathbf{U}^i) = \sum_{f=1}^{F}\log(V_f^i) + \sum_{f=1}^{F}\left(\frac{||P_f - U_f^i||^2}{\left(V_f^i\right)^2}\right)}$$

Clustering algorithm

Each location is a gaussian mixture distribution with
 k clusters with means and variances *Uk* and *Vk*
Wk the probability that an observed packet belongs to
 a particular cluster k.
Uk,*Vk* and *wk* are the three parameters.
Paremeters estimated using variational
 Bayesian inference.

Classification algorithm

Pinloc calculates macro location based on Wifi SSIDs
and shortlists the spots within this macro location.
Candidate set C
Define the distance between a given packet P and a spot Si as

$$\boxed{d(\mathbf{P}, S_i) = \min_{\mathbf{U}^i \in Z_i, AP(\mathbf{U}^i) = AP(\mathbf{P})} d(\mathbf{P}, \mathbf{U}^i)}$$

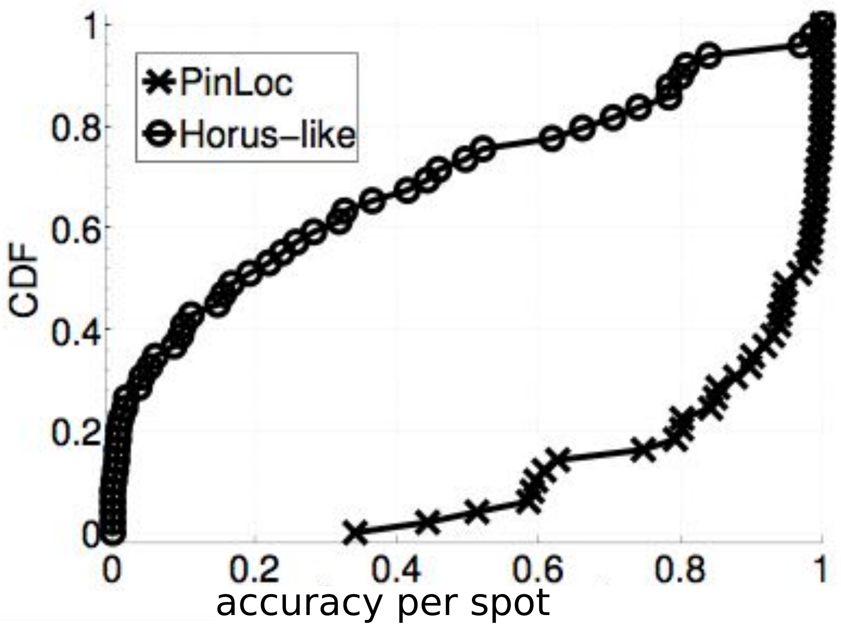
PinLoc Evaluation

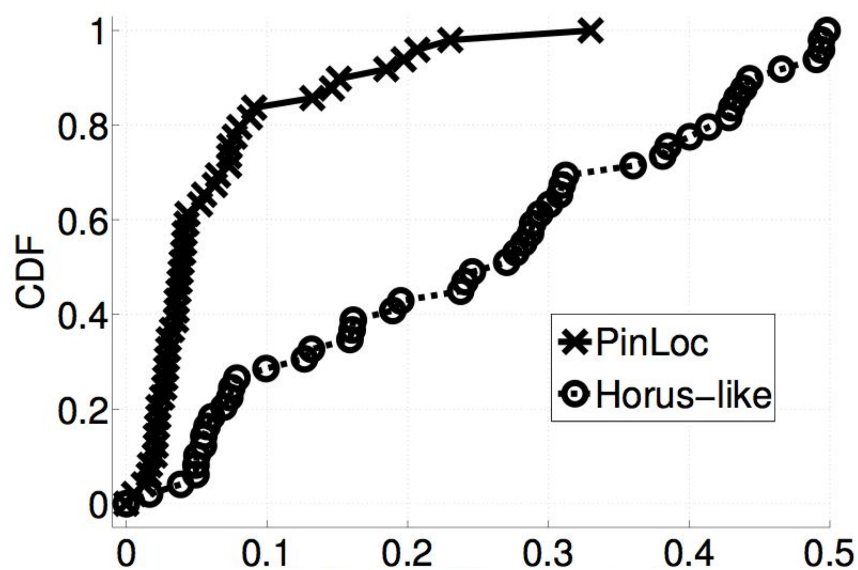
- Evaluated PinLoc (with existing building WiFi) at:
  - Duke museum
  - ECE building
  - Café (during lunch)

- Roomba calibrates
  - 4m each spot
  - Testing next day

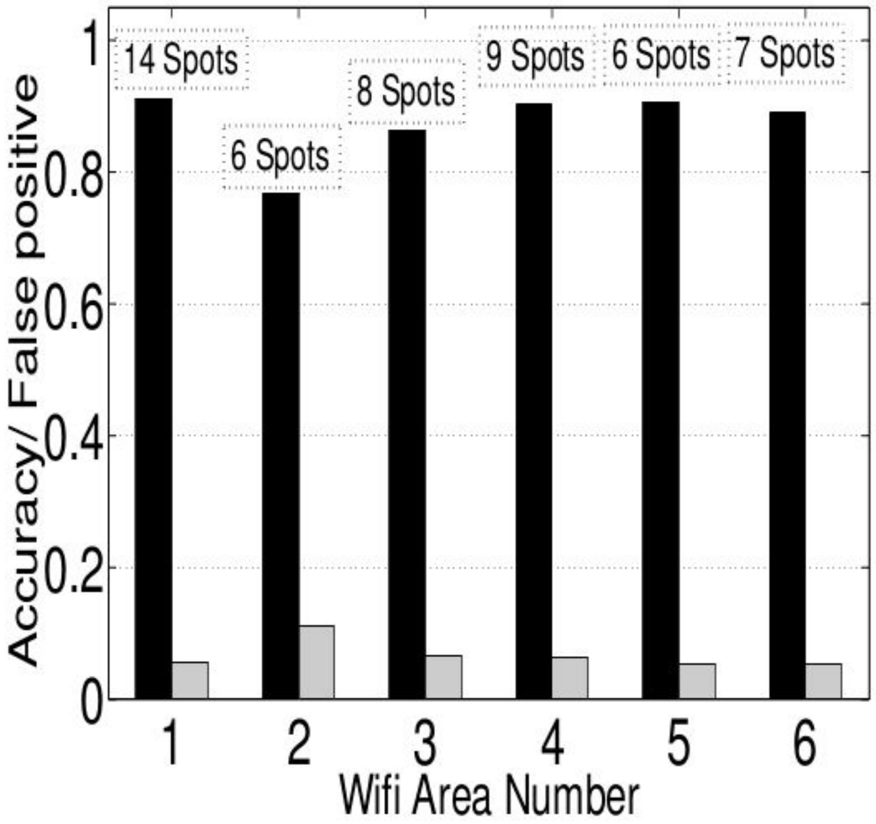
Performance

- 90% mean accuracy, 6% false positives
- WiFi RSSI is not rich enough, performs poorly $/ \quad 20\%$ accuracy





Accuracy on a per-Wifi macro location



Performance

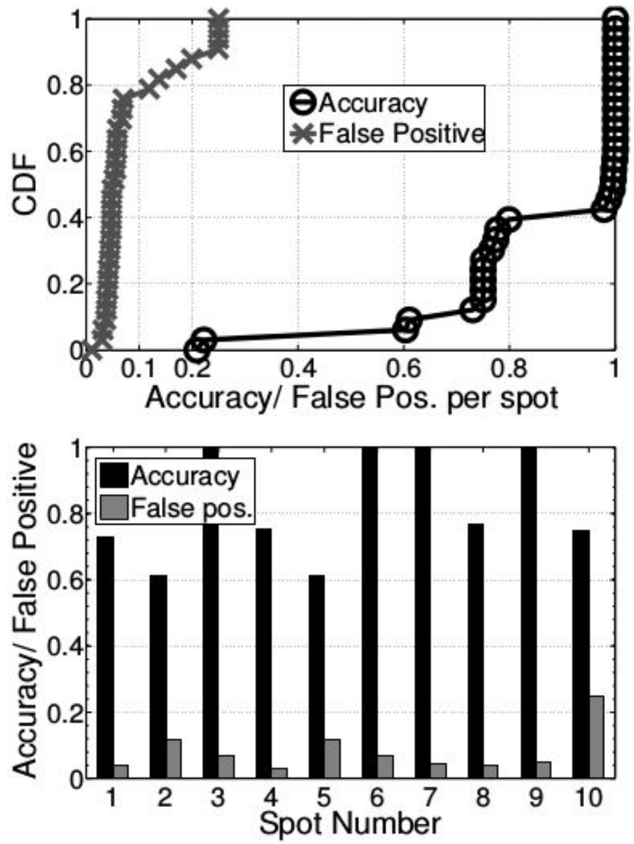


Figure 13: Pinloc performance in student center (a) Accuracy, false pos., (b) Performance of adjacent spots.

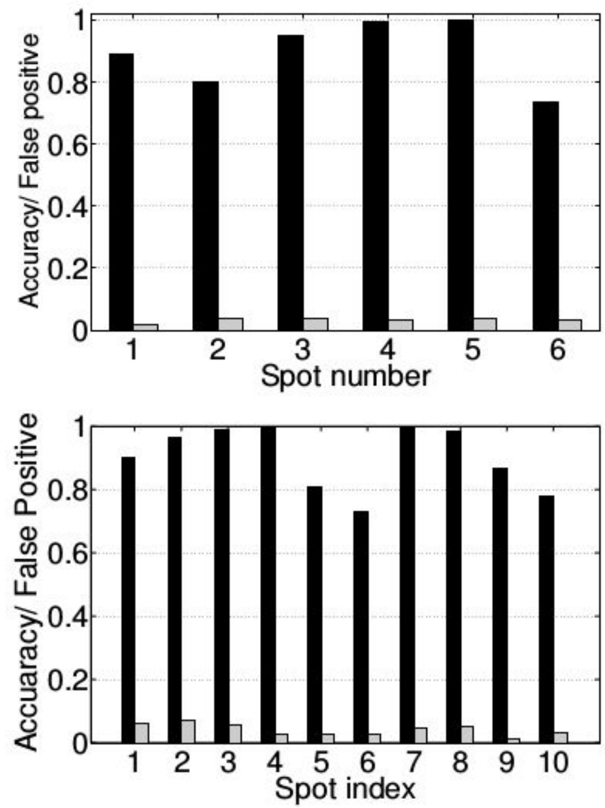
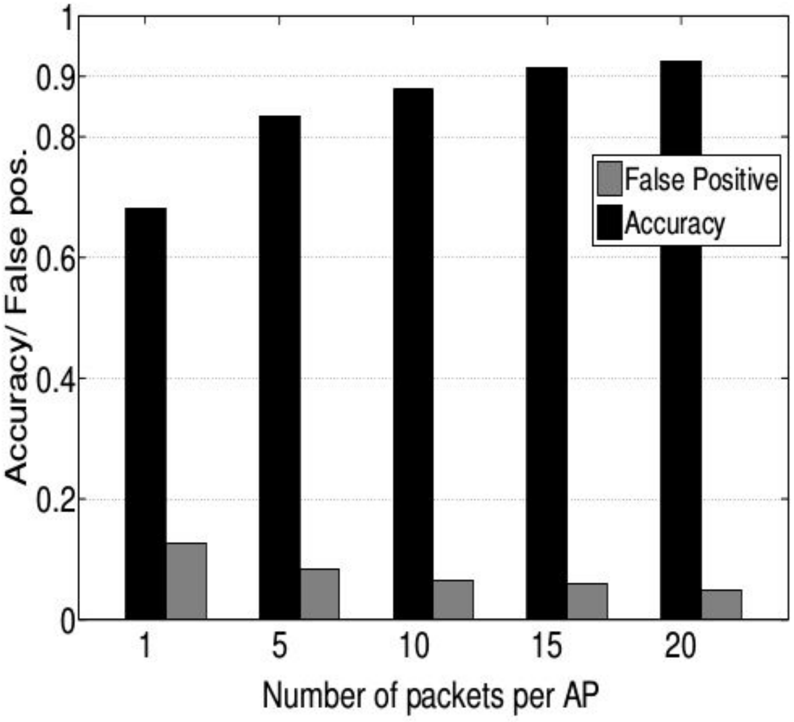


Figure 14: PinLoc performance in cafeteria and museum (a) Accuracy and FP per spot in cafeteria. (b) Accuracy and FP per-spot in the museum.

Impact of Parameters

- number of test packets
- number of Aps
- war-driving
- mobility
- old training data

Impact of number of test packets

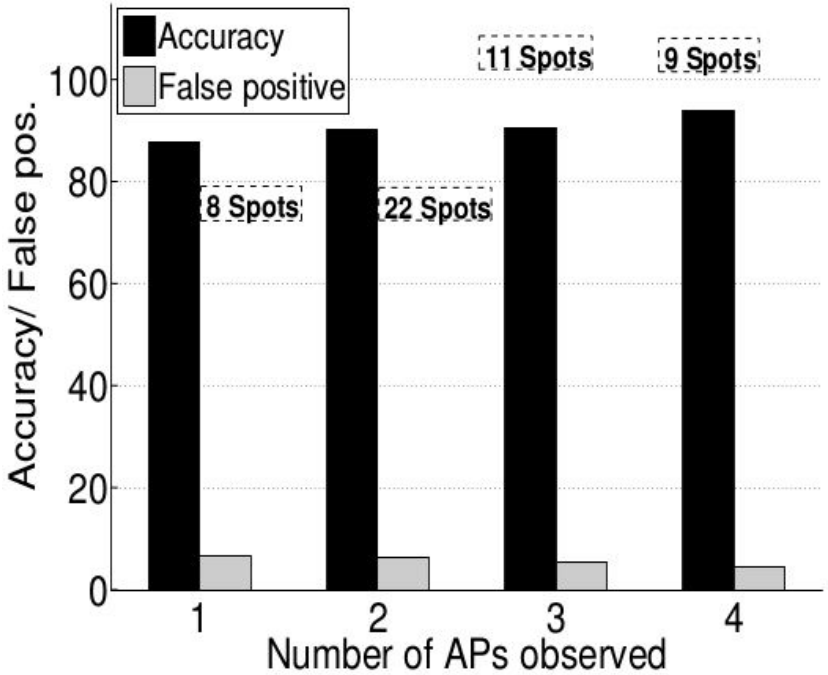


With 10 packets per AP, mean accuracy
is 89% (7% false positives)

With 1 packet the mean accuracy reduces
to 68%(14% false positives)

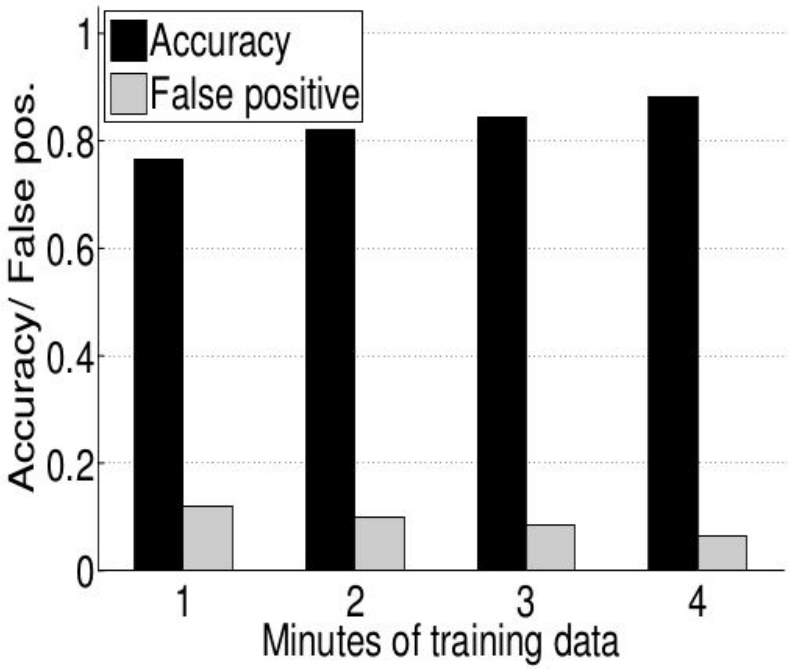
Single reading may randomly match
with an incorrect spot.

Impact of the number of APs



Even with single AP visible the mean
 accuracy is over 85%
 (below 7% false positives)
Significant improvement as other Wi-fi
 based localization method need at
 least 3 Aps.

Impact of war-driving



Short wardriving records fewer CFRs incurring the possibility of overlooking important ones.

Reasonable performance observed

even for 1 minute of wardriving

Impact of mobility

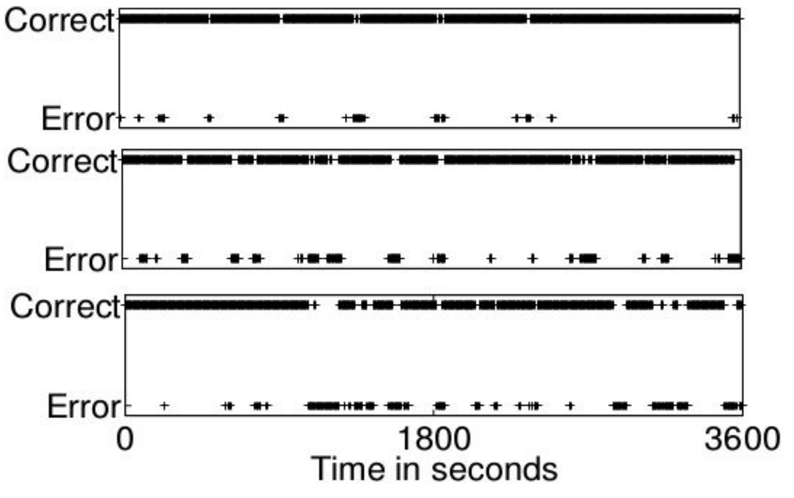


Figure 16: Success of PinLoc localization over time for three spots and over an interval of 1 hour.

Cafeteria scenerio
Time interval – 1hr
Mean accuracy – 85% (7% false positives)
Time instants of failure are short and
evenly distributed.

Impact of old training data

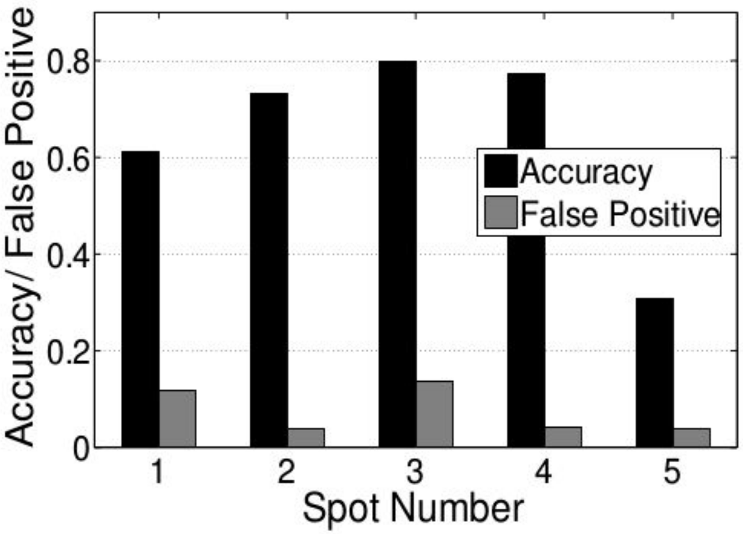


Figure 17: Accuracy of 5 spots tested 7 months after training.

Need fresh rounds of wardriving for
spots affected by significant
environmental changes.

With 5 spots observed after 7 months
median accuracy of 73% found