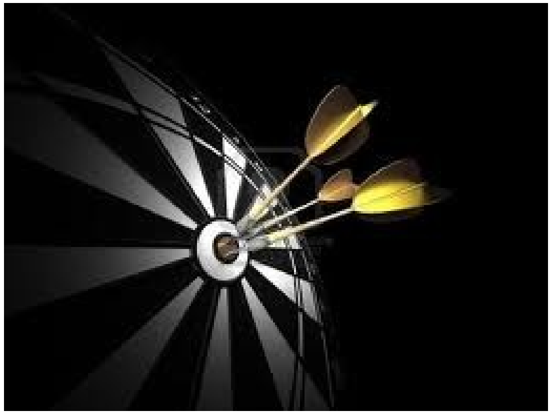
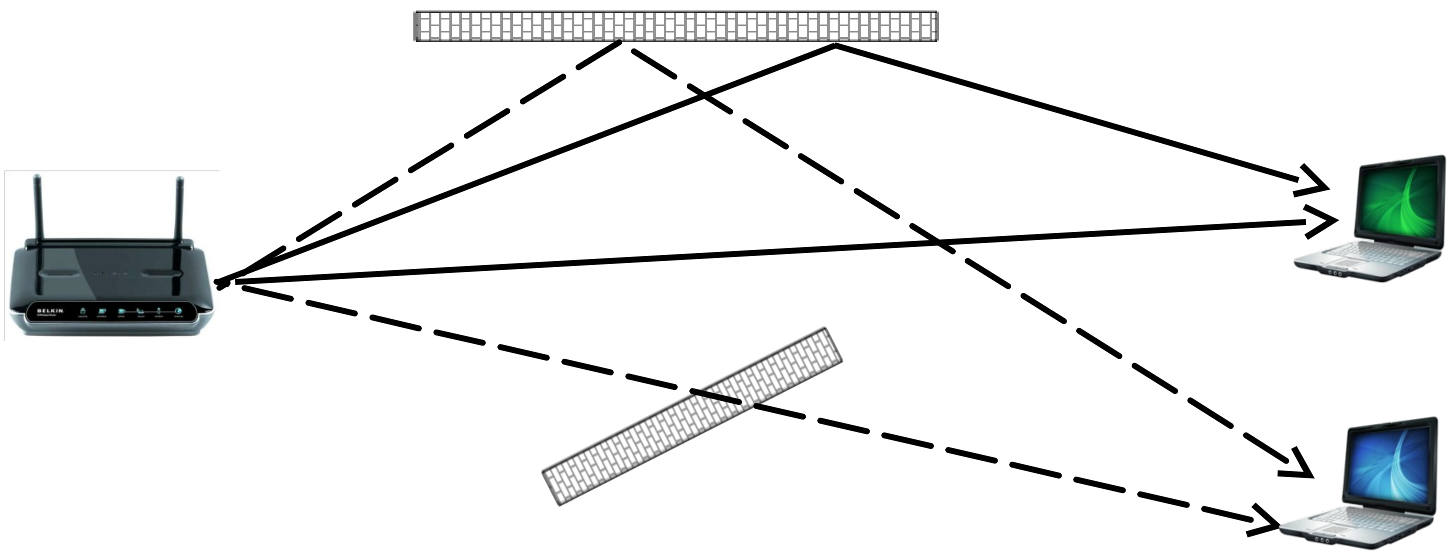


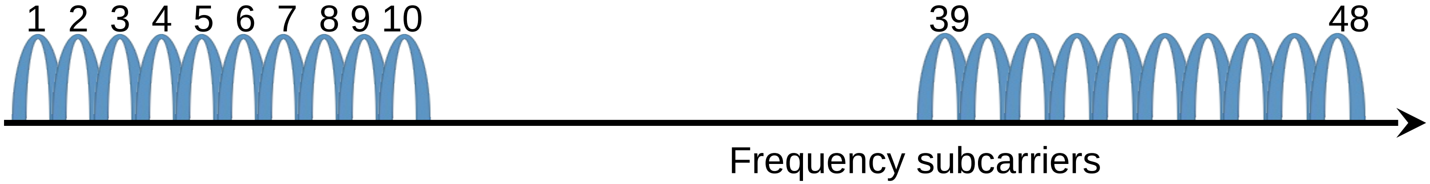
PinLoc*

*Planned for deployment in Duke's Nasher Art Museum

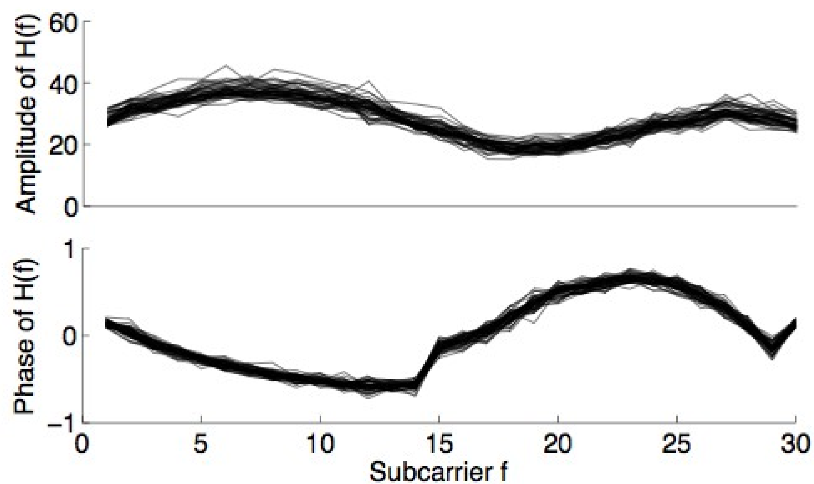


Fingerprinting Wireless Channel

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



  - Intel card ponse per s ier



Is WiFi Channel Amenable to Localization?

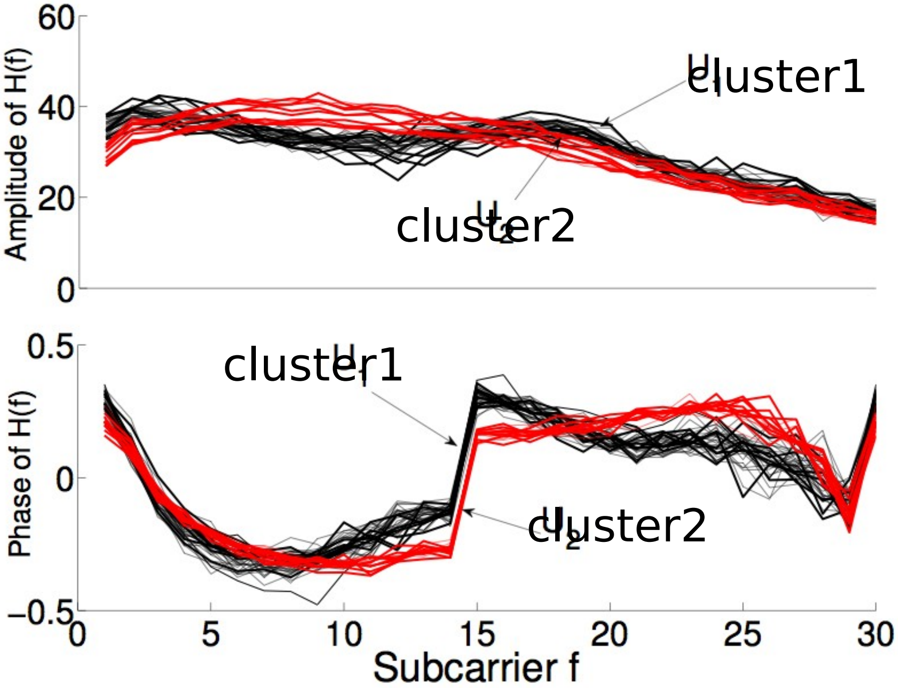
- Two key hypotheses need to hold:

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**

Variation over Time

- Measured channel response at different times
  - Using

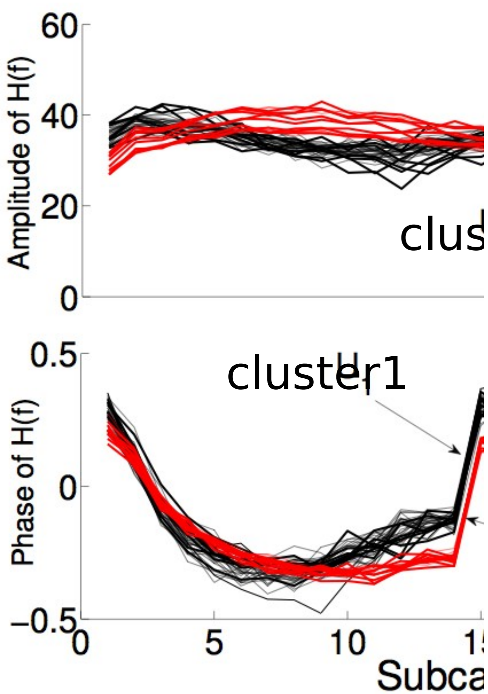


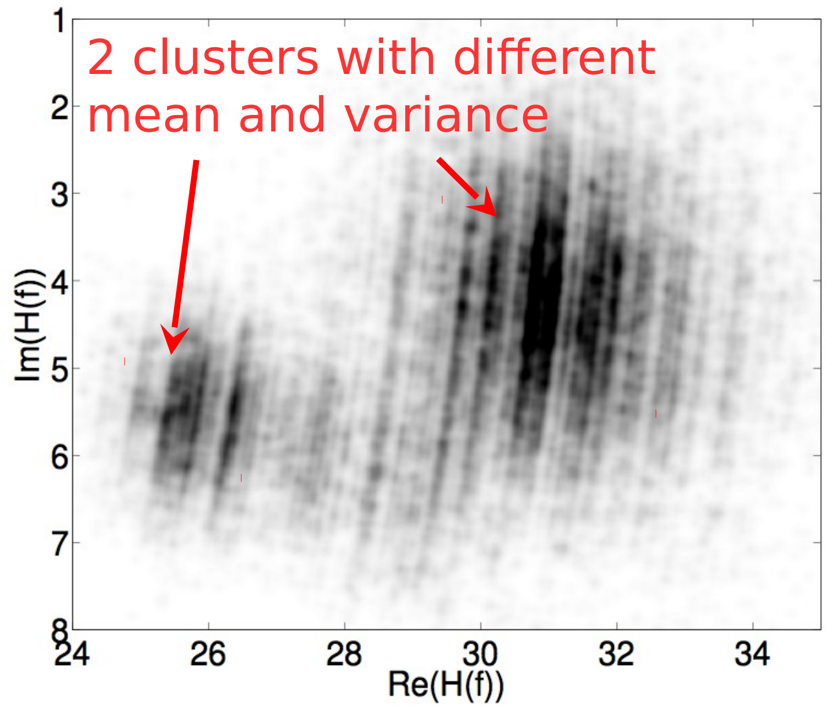
Observe: Frequency responses often clustered at a location

But not necessarily one cluster per location

Variation over Time

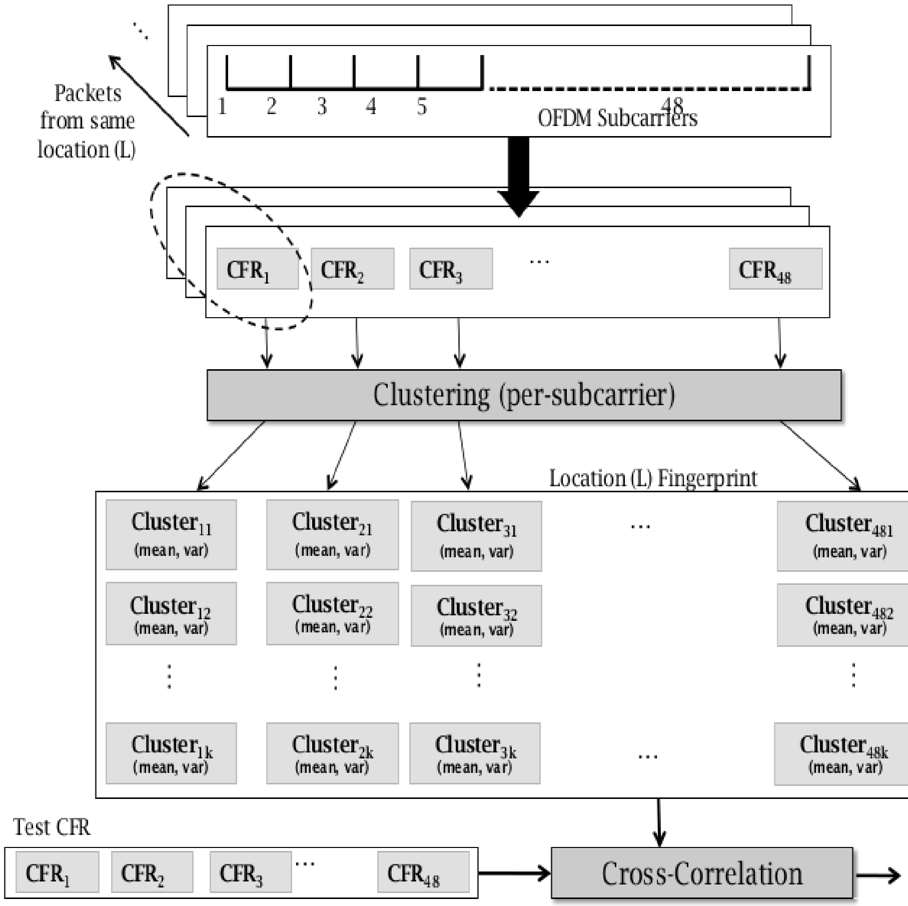
- Measured channel response at different times
  - Using Intel cards





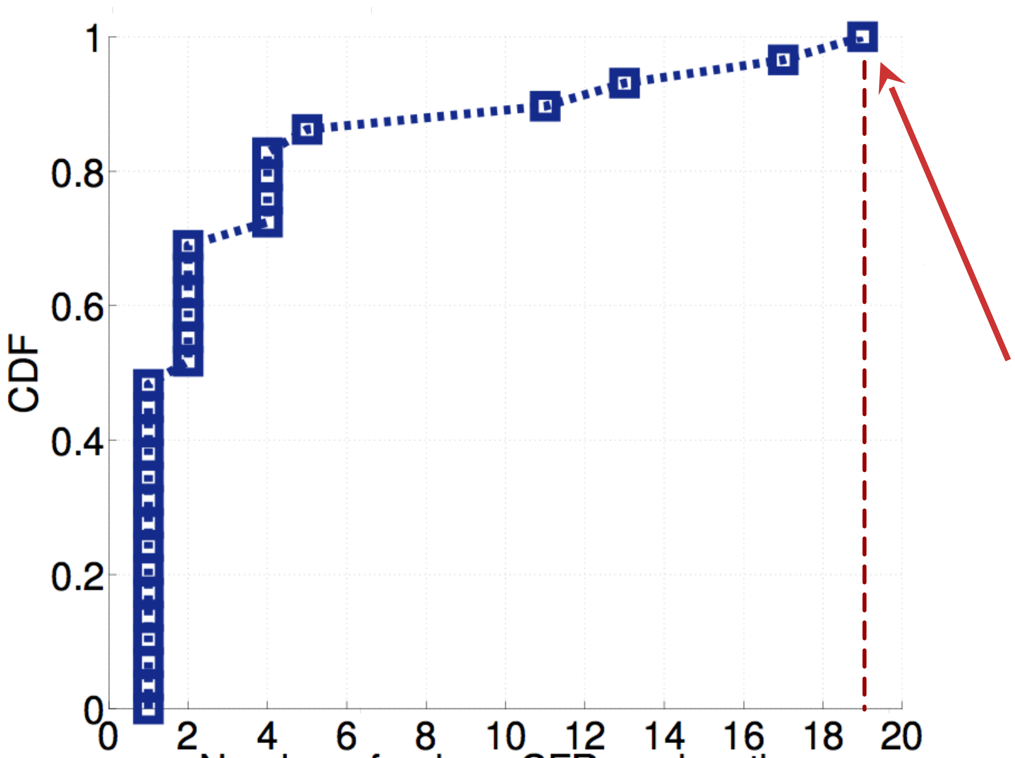
But not necessarily one cluster per location

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

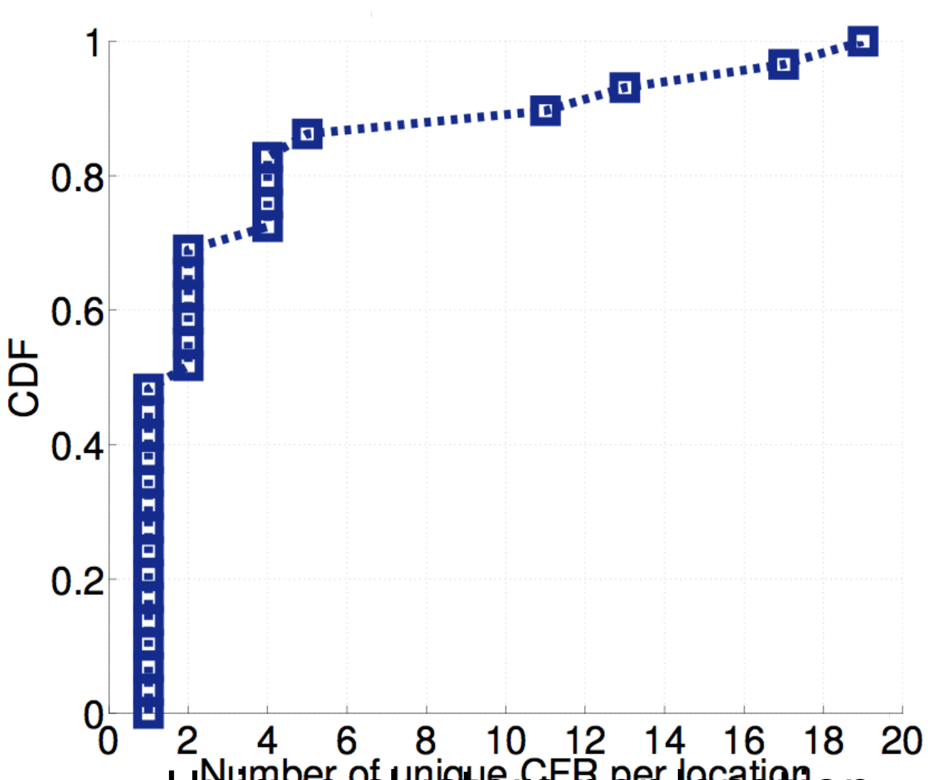
How Many Clusters per Location?



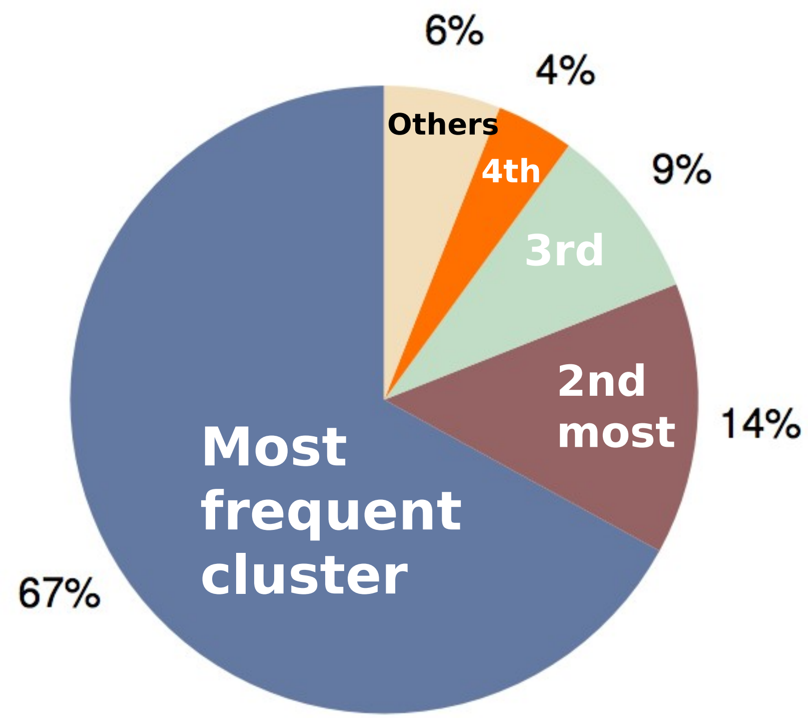
Unique clusters per location

Do all 19 clusters occur
with same frequency?

Cluster Occurrence Frequency



Unique clusters per location

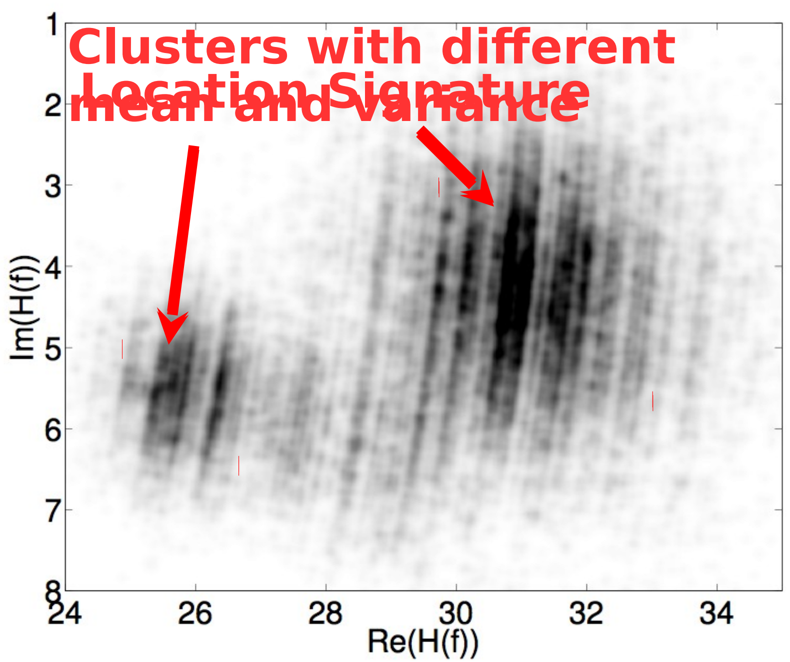


3 to 4 clusters heavily dominate, need to learn these signatures

Is WiFi Channel Amenable to Localization?

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

2. **Spatial**
   - **Channel** … **ed to be differen**…

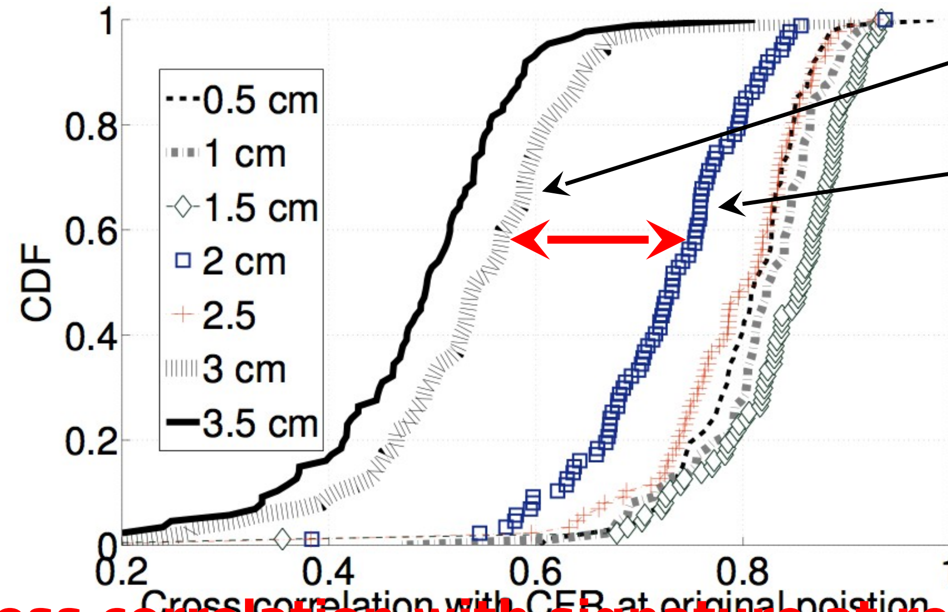


What is the Size of a Location?

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

What is the Size of a Location?

- Localization granularity depends on size
  - several meters



3 cm apart

2 cm apart

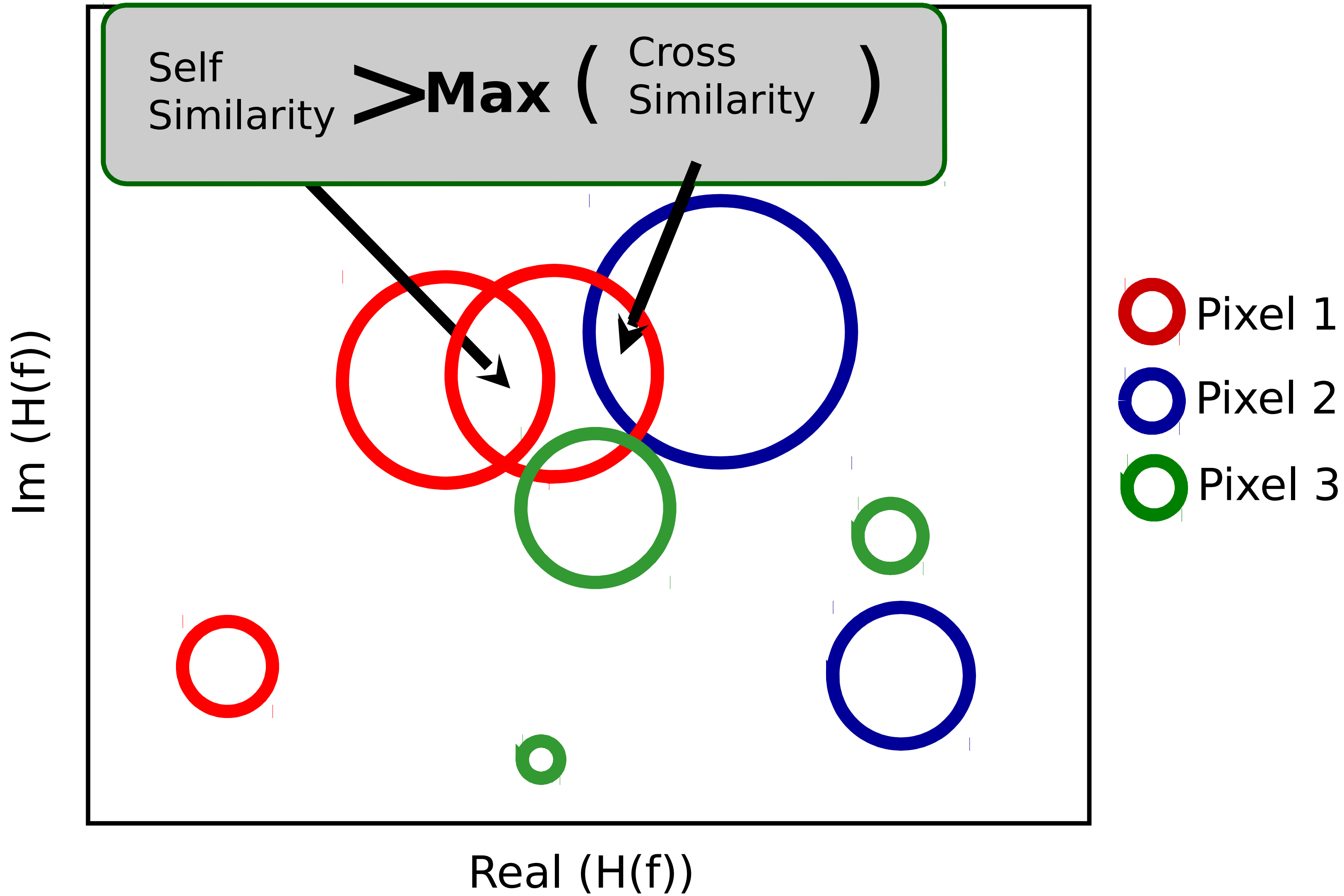
Cross correlation with signature at reference location

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

ation” as 2cm x 2cm area- call them pixels

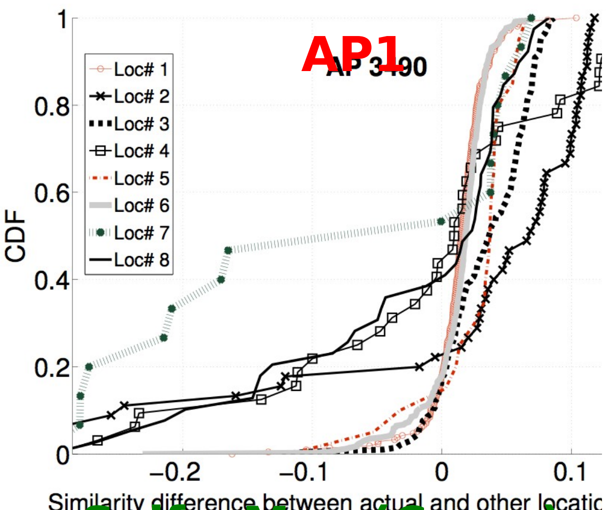
ance changes every 2-

But … Will all pixels have unique signatures?

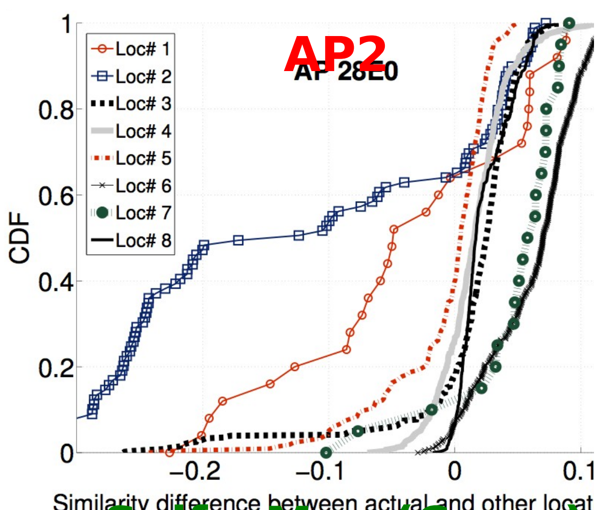


For correct pixel localization:

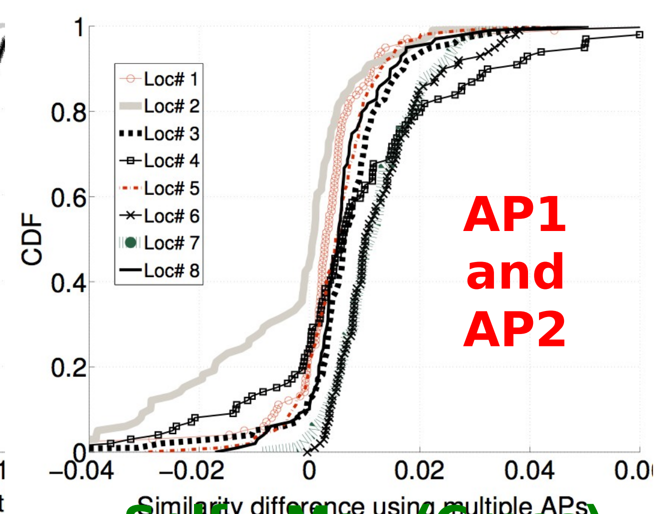
Self Similarity − **Max** (Cross Similarity) > 0



Self – Max (Cross)



Self – Max (Cross)



Self – Max (Cross)

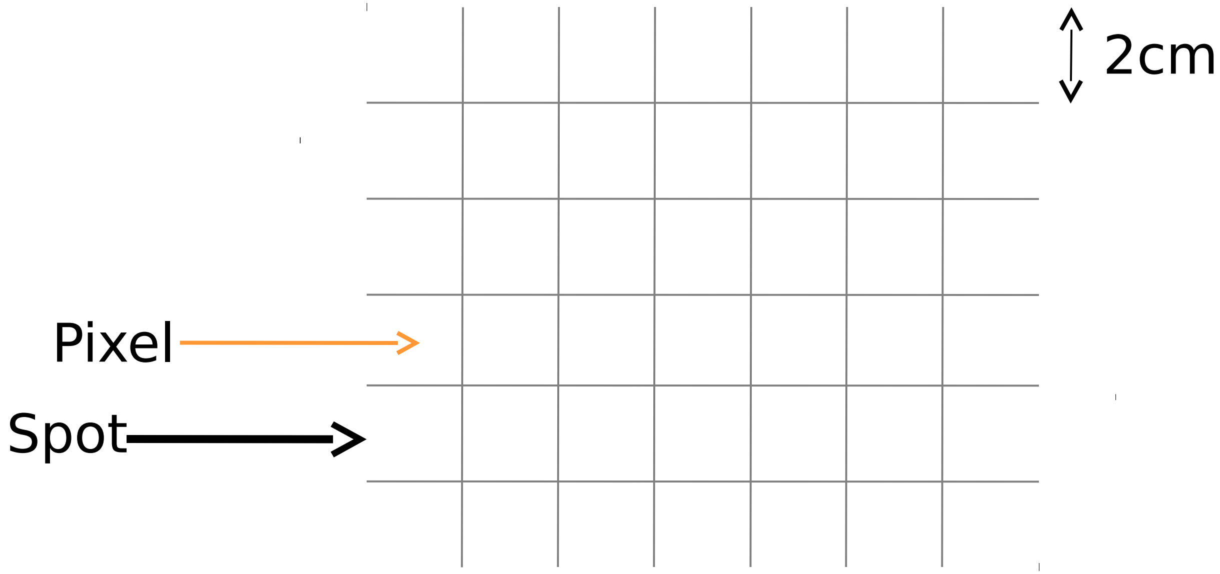
67% pixel accuracy even with multiple APs

67% accuracy inadequate …
can we improve accuracy?

Opportunity:

- Humans exhibit natural (micro) movements
- Likely to hit several nearby pixels
- Combine pixel fingerprints into super-fingerprint

From Pixels to Spots



Combine pixel fingerprints from a 1m x 1m box.

Intuition: low probability that a set of pixels will all match well with an incorrect spot

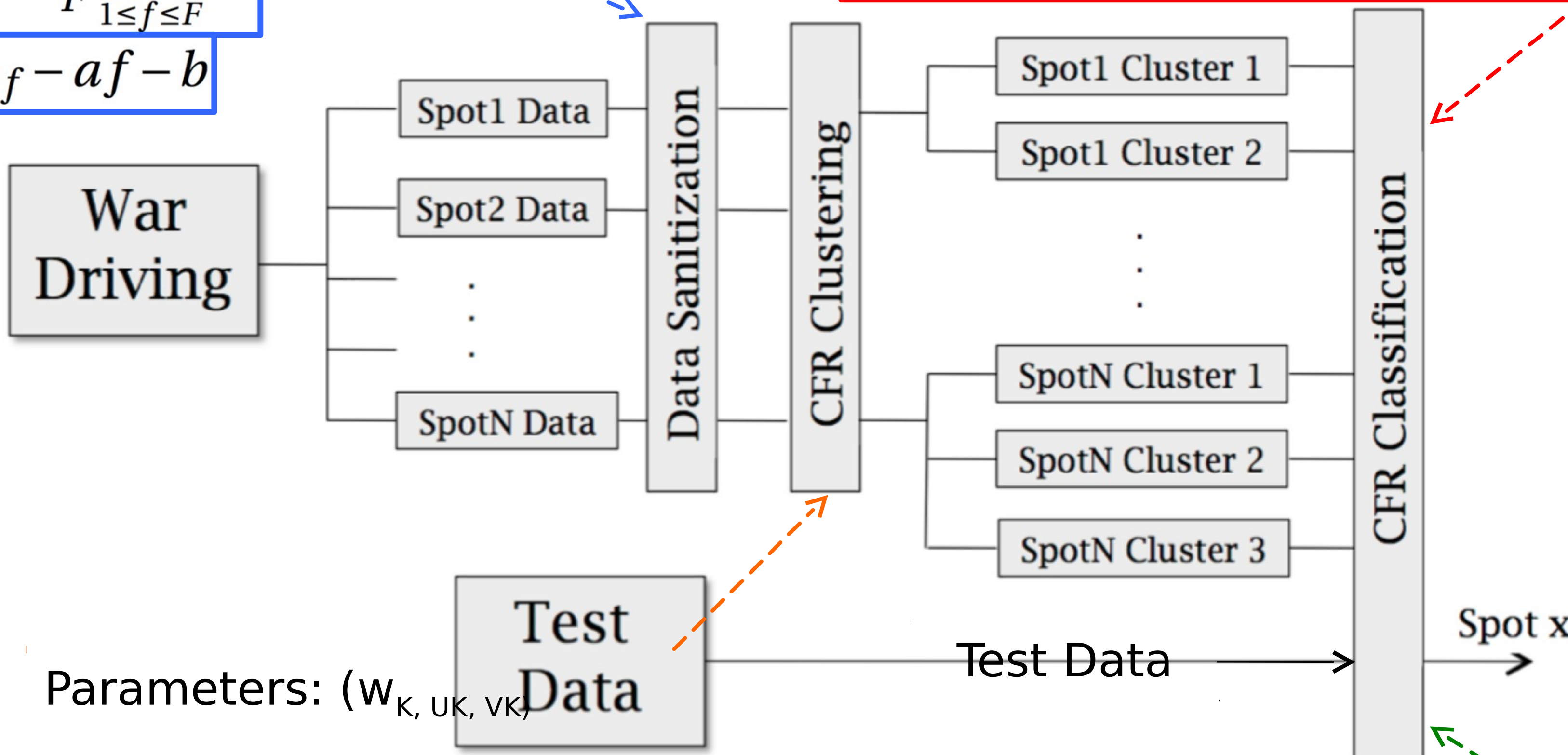
PinLoc: Architecture and Modell

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

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

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

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



Parameters: $(w_{K,\,UK,\,VK})$

Variational Inference (Infer.NET)

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

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) a
  - –Duke museum
  - –ECE building
  - –Café (during lu

- Ro bra
  - –4 t
  - –Te day

Performance

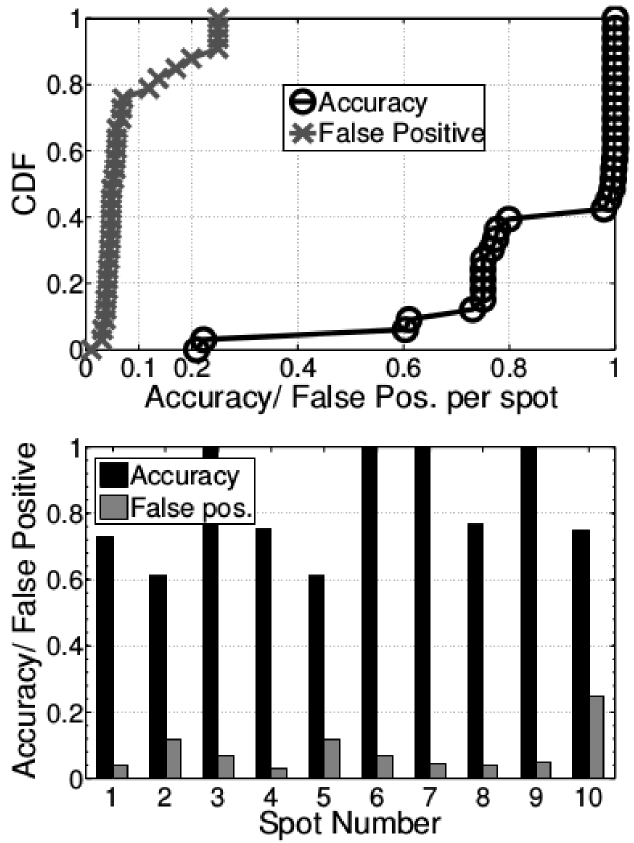


Figure 13: Pinloc performance in student center (a) Accuracy, false pos., (b) Performance of adjacent spots.

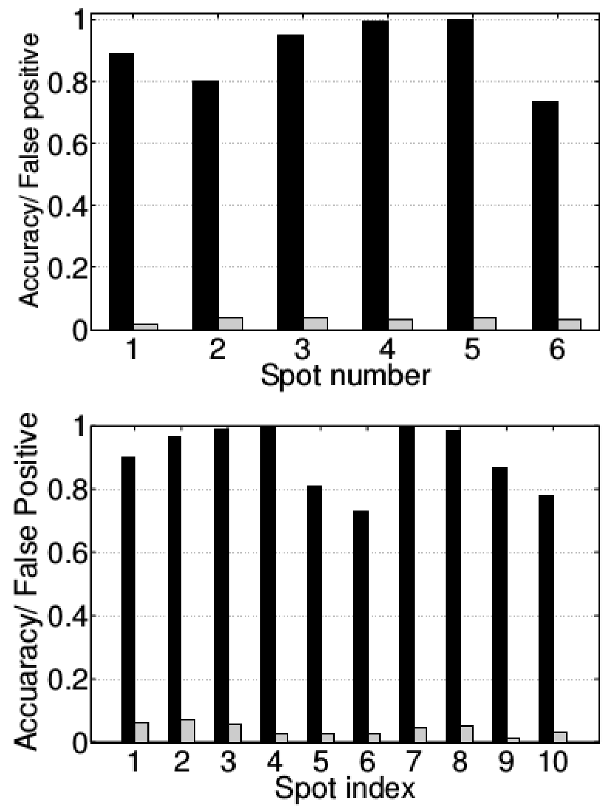
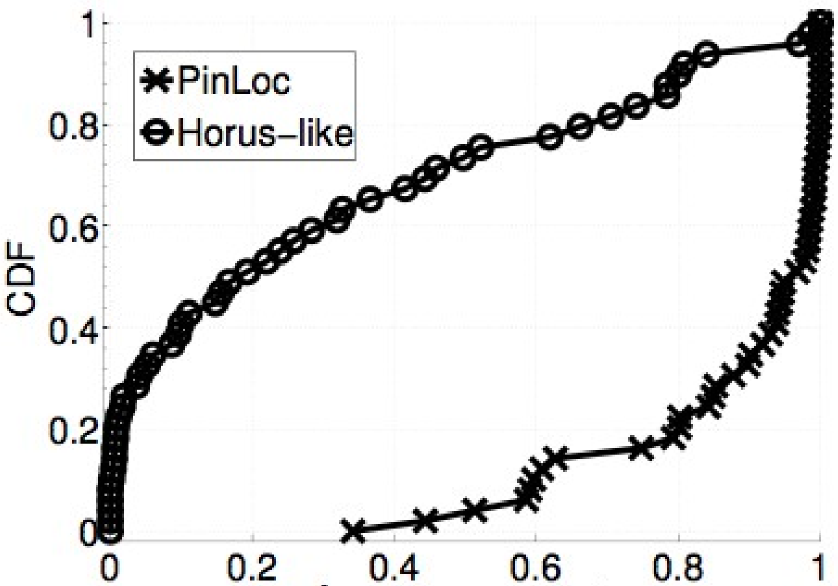


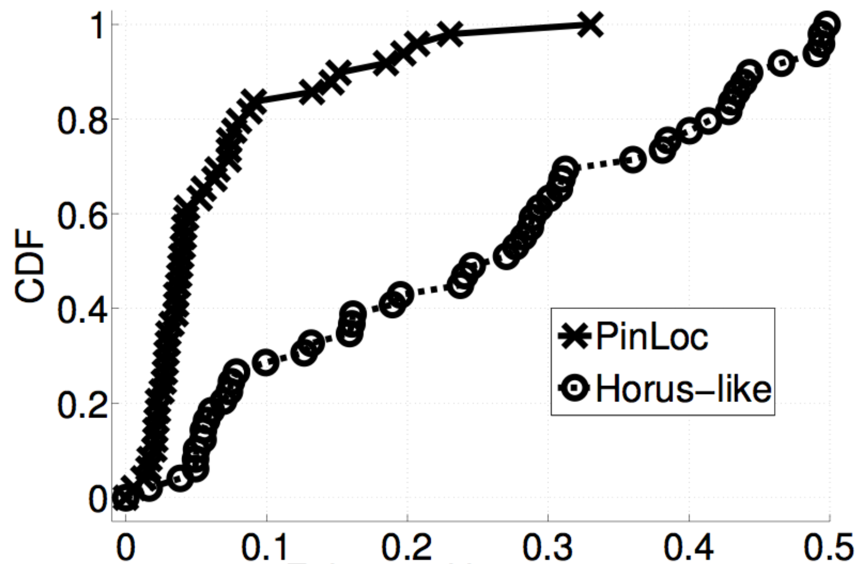
Figure 14: PinLoc performance in cafeteria and museum (a) Accuracy and FP per spot in cafeteria. (b) Accuracy and FP per-spot in the museum.

Performance

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



Accuracy per spot

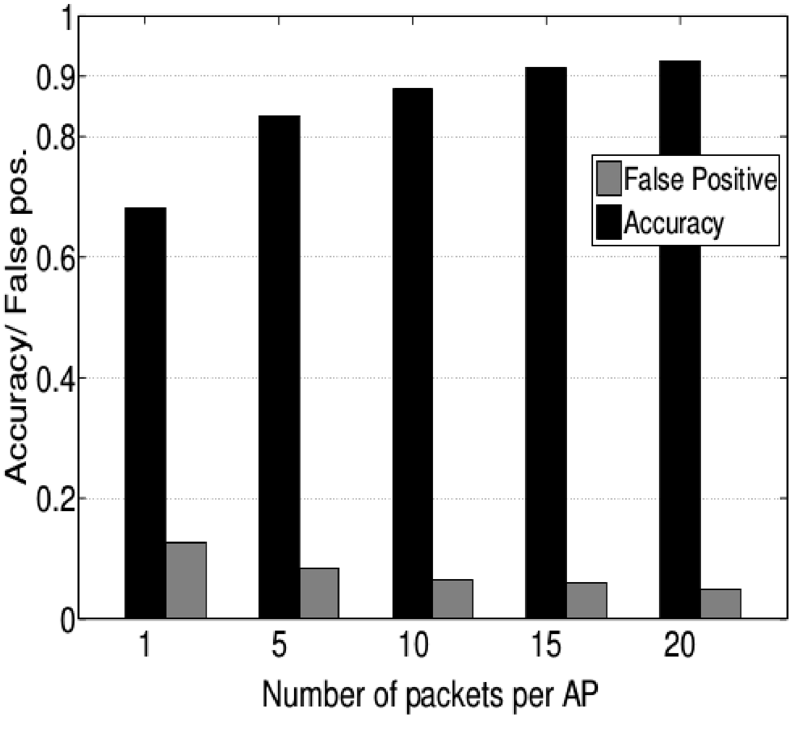


False positive per spot

Impact of Parameters

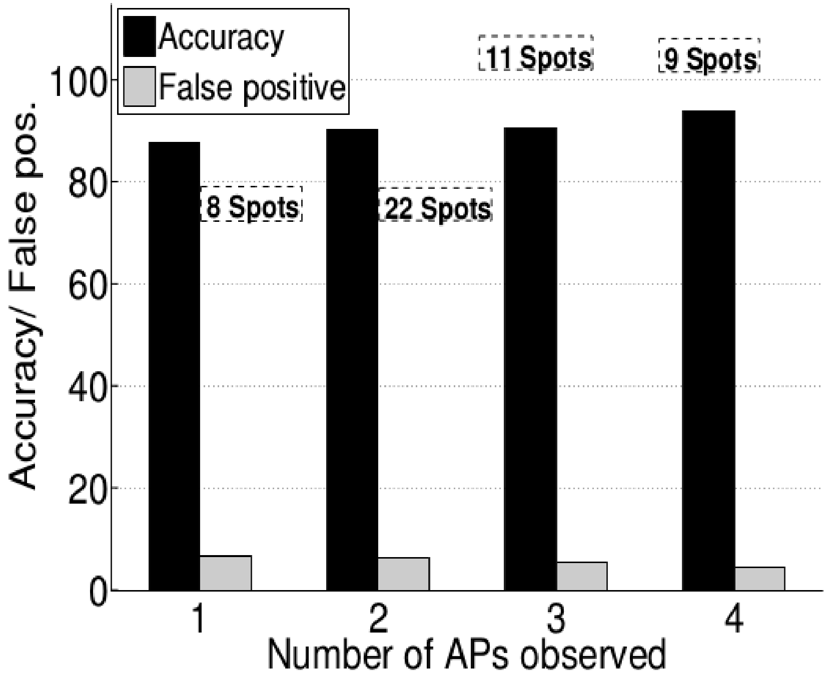
- number of test packets
- number of Aps
- war-driving
- mobility
- old training data

Impact of number of test packets



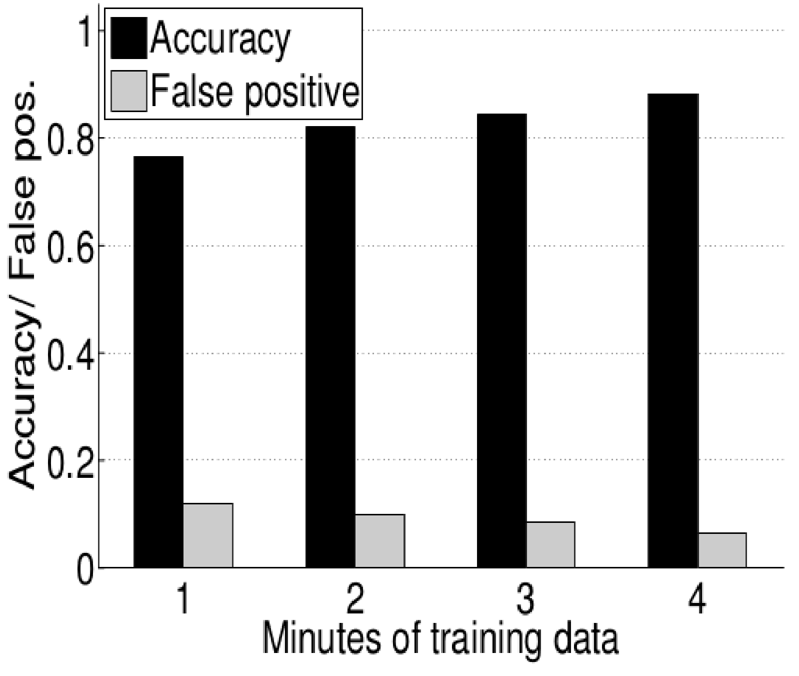
- With 10 packets per AP, mean accuracy is 89% (7% false positives)
- With 1 packet the mean accuracy reduces to 68%(14% false positives)
- Single reading may randomly match with an incorrect spot.

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

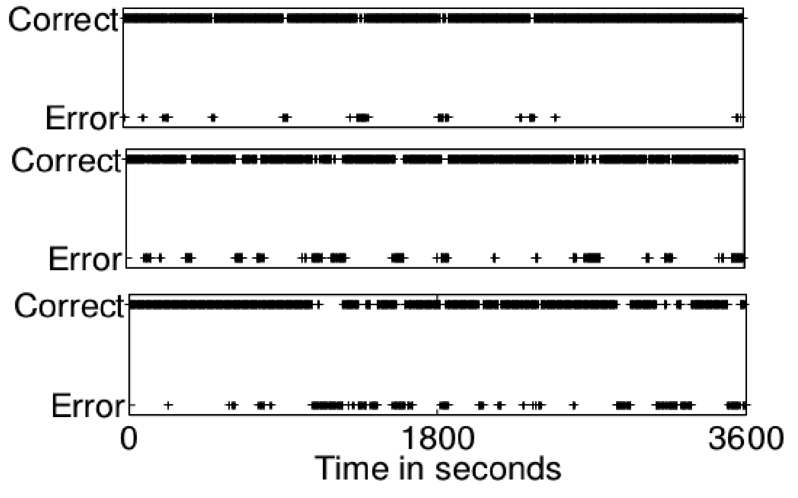


Figure 16: Success of PinLoc localization over time for three spots and over an interval of 1 hour.

- 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

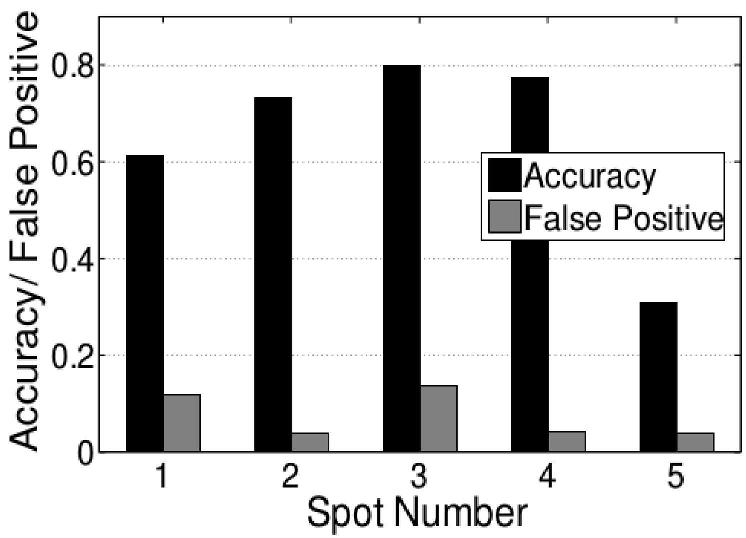


Figure 17: Accuracy of 5 spots tested 7 months after training.

- Need fresh rounds of wardriving for spots affected by significant environmental changes.
- With 5 spots observed after 7 months median accuracy of 73% found