Indoor Localization: WiFi Fingerprinting

Amitangshu Pal

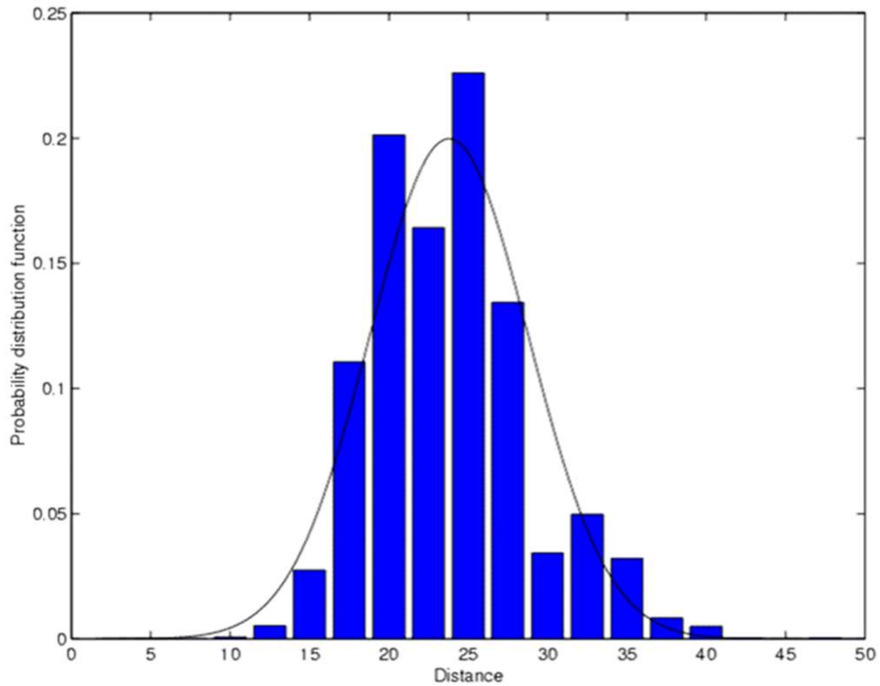
Indoor Localization

https://www.geospatialworld.net/blogs/indoor-positioning-indoors-gps-stops-working/

Indoor Localization

- Why GPS localization is not used in indoors?
  - GPS cannot work indoors
  - GPS power consumption is very high

- Why not use WiFi APs as satellites and use trilateration?
  - Heavy multipaths in indoor environment
  - WiFi APs/routers are not precisely clock synchronized

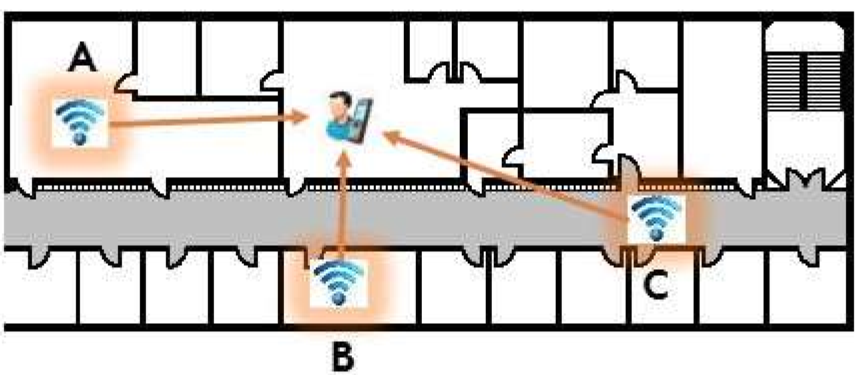


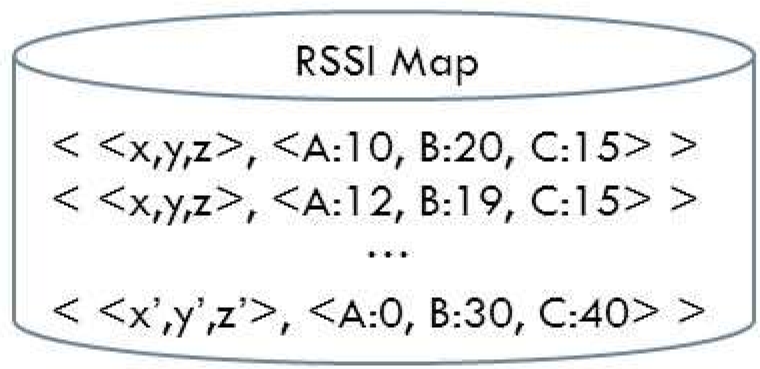
Indoor Localization: Deterministic Approach

RADAR: Offline Phase

- WiFi fingerprinting → Offline phase and Online phase
- For every location and for every orientation of these locations, measure:
  - $\ll x, y, z > < RSSI^A, RSSI^B, RSSI^c \gg$

- RSSI values are averaged over multiple measurements

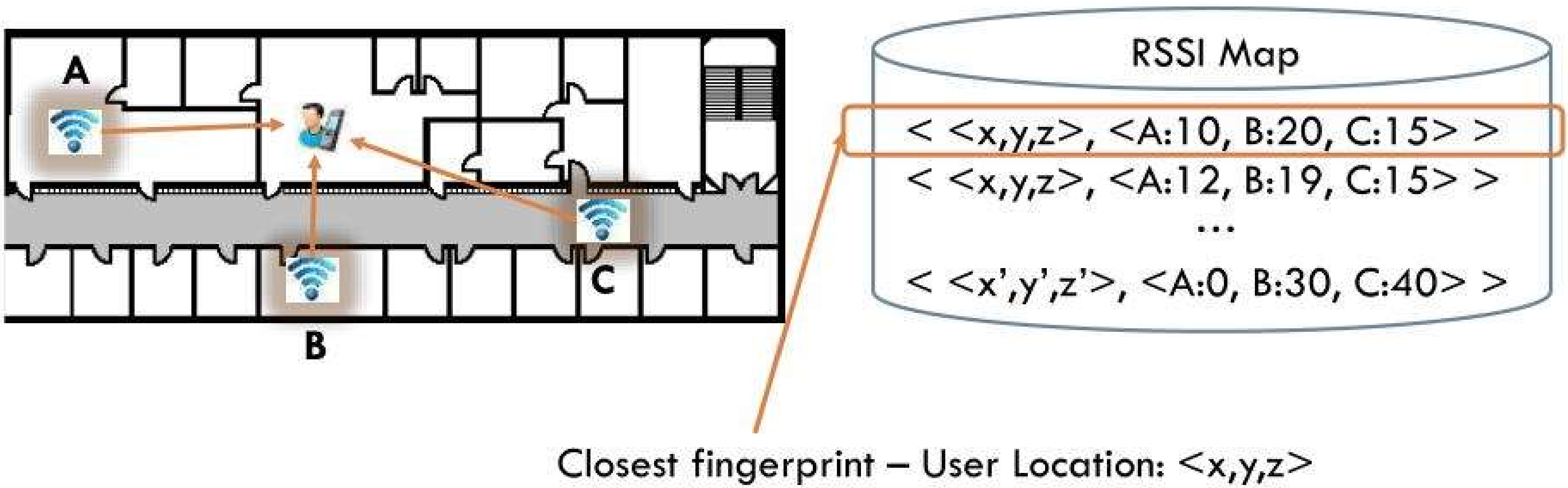




https://www.slideserve.com/holly/indoor-localization-using-fingerprinting

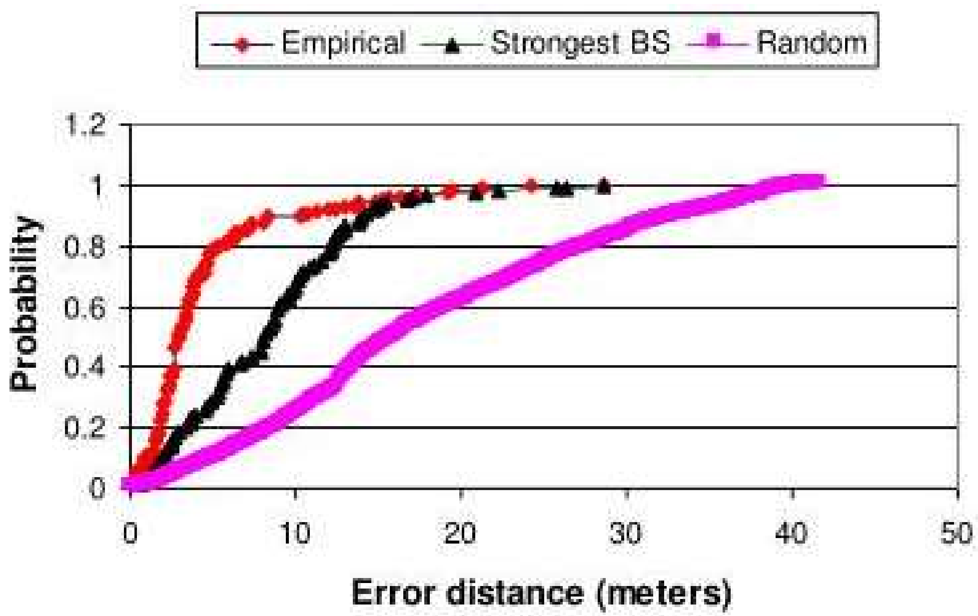
RADAR: Online Phase

- At any target location, record all the RSSI values:
  - $\ll RSSI^A, RSSI^B, RSSI^c = < A: 11, B: 20, C: 15 \gg$
- Find the location $< x, y, z >$ that has the closest fingerprint (or **nearest neighbor**) in the RSSI map
  - $< x, y, z >$ then becomes the location of the user

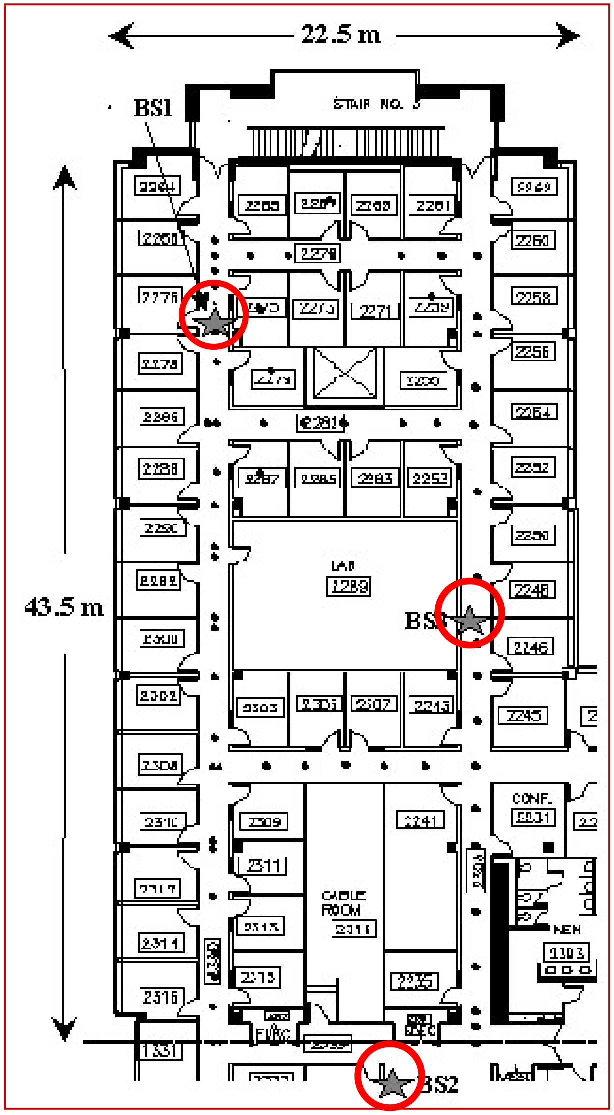


RADAR: Experimental results

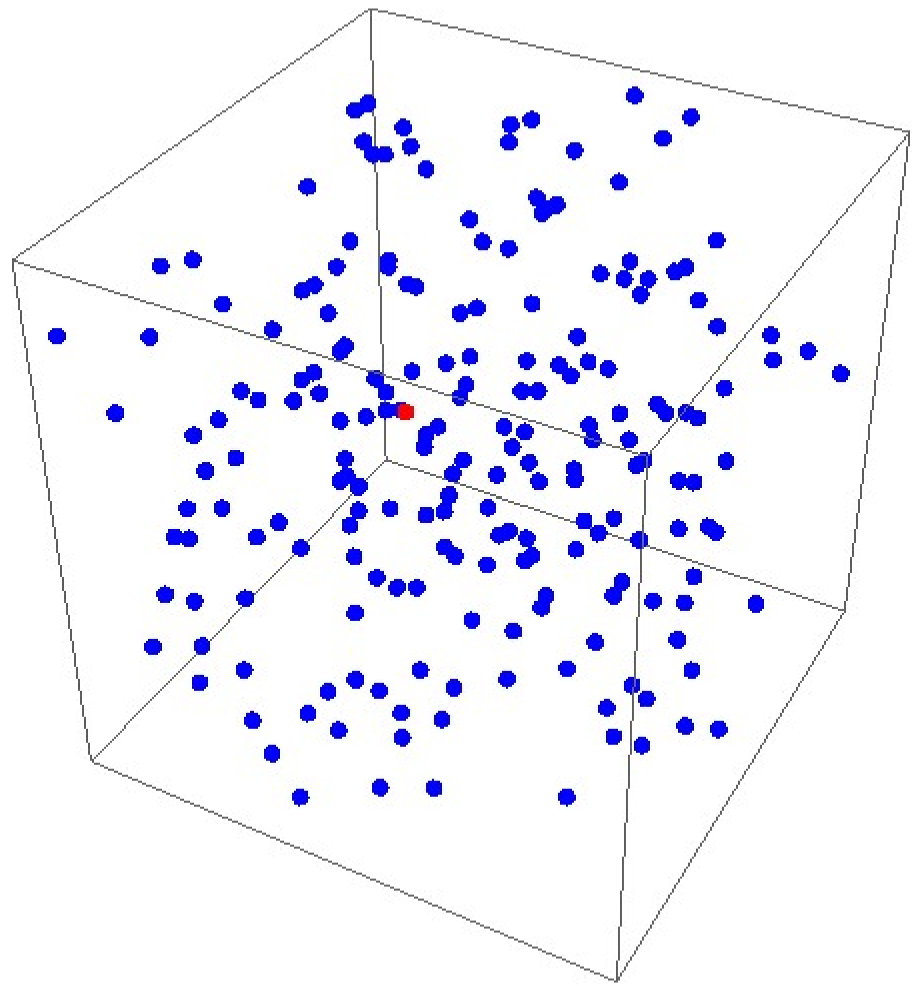
- Floor layout:
  - Black Dots: locations where empirical signal strength info was collected
  - Large Stars: Access points



Median error: 2.94 meters
90% error: 10 meters

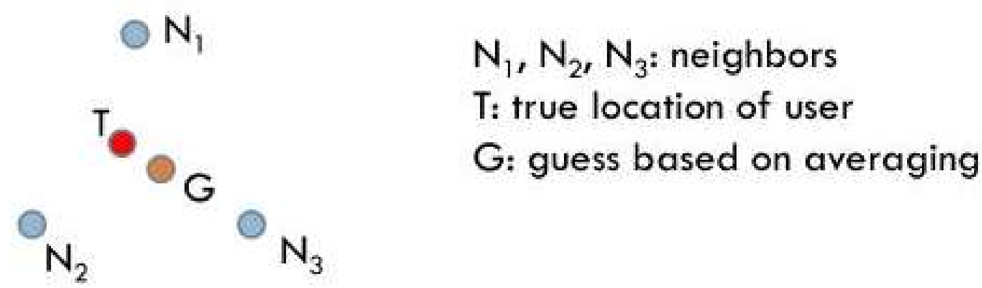


RADAR: Experimental results

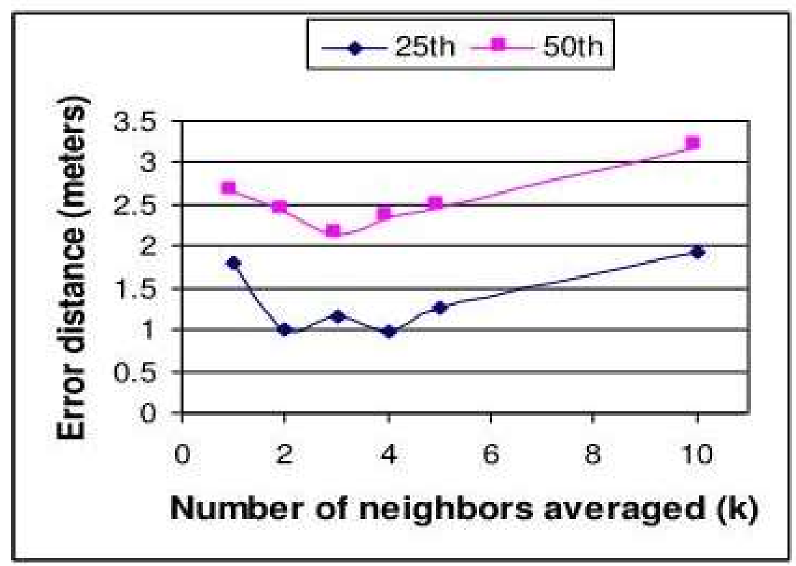


RADAR: Experimental results

- Lets not limit to just nearest data point
- Find ***k*-nearest neighbours**:
  - Finding the right k is challenging



Median error with $k = 3 \rightarrow$ 2.13 meters



https://cupdf.com/document/introduction-to-the-global-positioning-system-56bc80026faa1.html

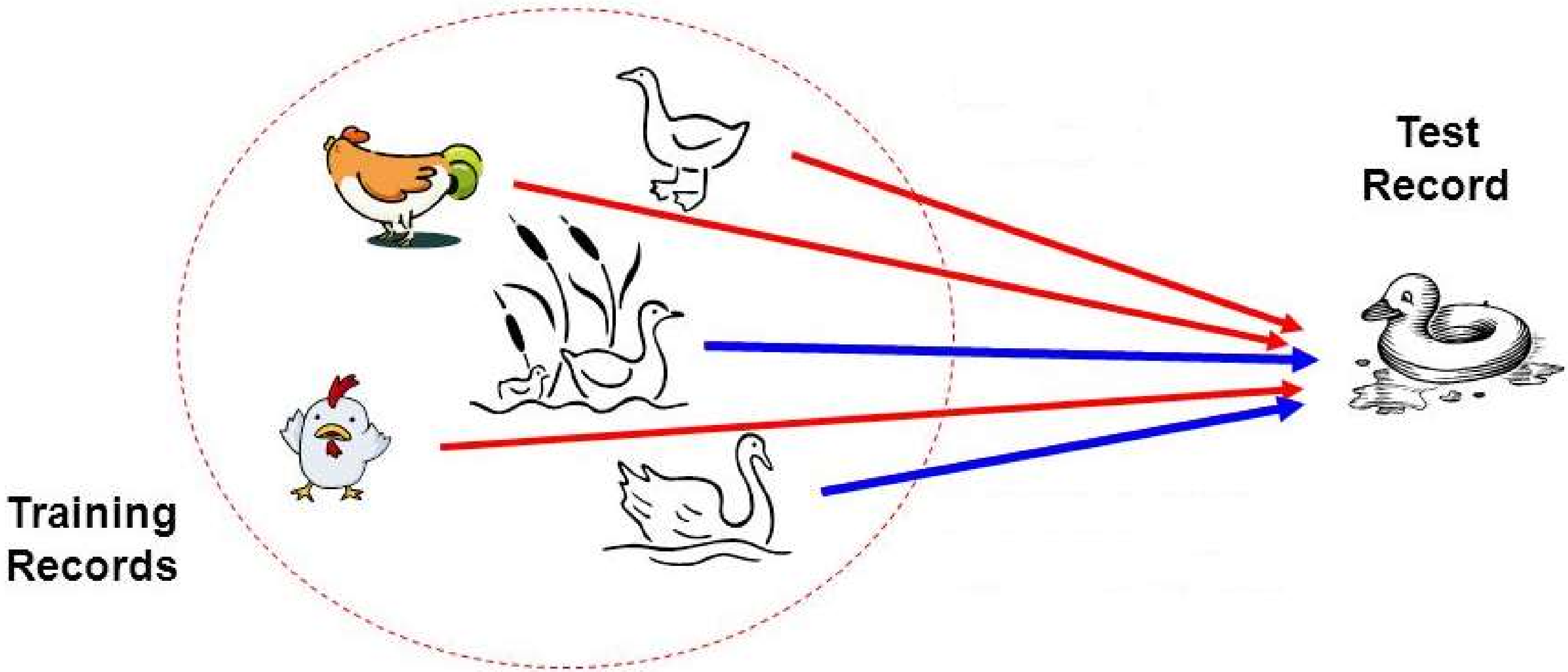
RADAR: Limitations

- Long time to gather all the empirical data
  - 1 floor= (70 locations) · (4 directions) · (20 samples)
  - No one wants to collect all that data for a whole office building
- If the access point moves, have to recollect all the data

Indoor Localization: Probabilistic Approach

Bayesian Classifier

- Principle
  - If it walks like a duck, quacks like a duck, then it is probably a duck



Src: D. Samanta

Bayesian Classifier

- Suppose, *Y* is a class variable and *X* = $\{X_1, X_2, \ldots\ldots, X_n\}$ is a set of attributes, with instance of *Y*.

INPUT (X)	CLASS(Y)
… … …	
… … …	…
$x_1, x_2, \ldots, x_n$	y_i
… … …	…

- The classification problem, then can be expressed as the class-conditional probability

$$P\big(Y = y_i | (X_1 = x_1) \text{ AND } (X_2 = x_2) \text{ AND } \ldots\ldots (X_n = x_n)\big)$$

Bayesian Classifier

- Bayesian classifier calculate this posterior probability using Bayes' theorem, which is as follows.
- From Bayes' theorem on conditional probability, we have

$$P(Y|X) = \frac{P(X|Y) \cdot P(Y)}{P(X)}$$

Note:

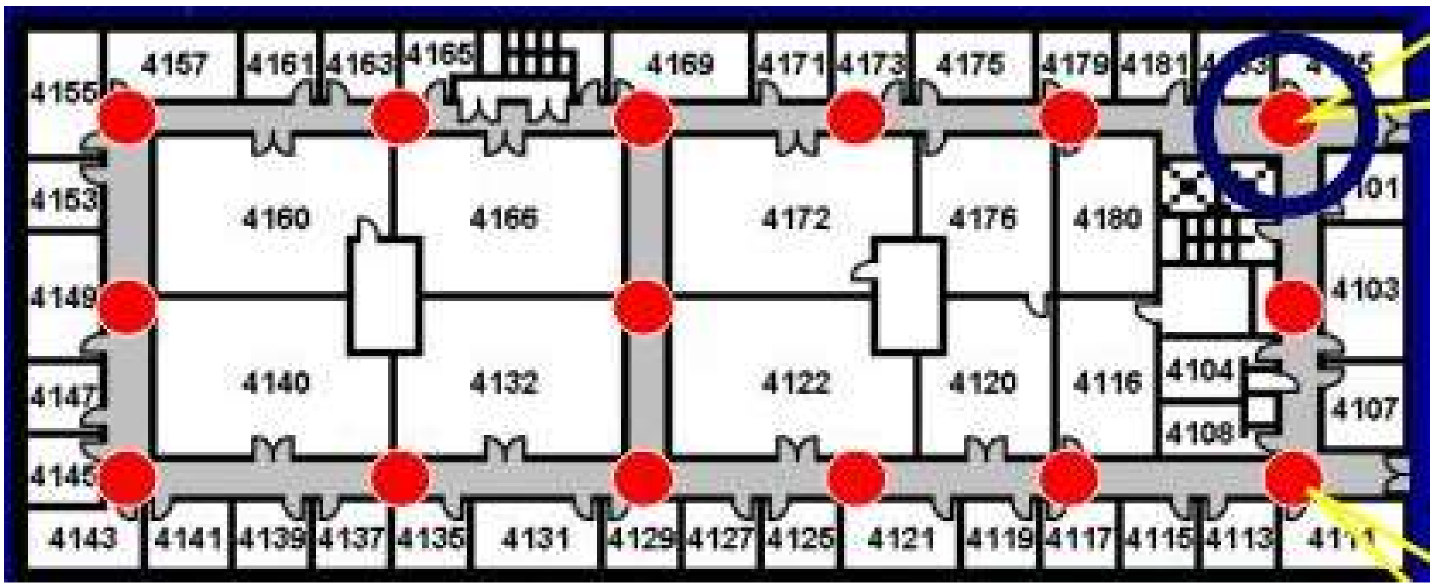
- $P(X)$ is called the evidence (also the total probability) and it is a constant.
- The probability *P(Y|X)* (also called class conditional probability) is therefore proportional to *P(X|Y)*$\cdot$ $P(Y)$.
- Thus, *P(Y|X)* can be taken as a measure of *Y* given that *X*.

$$P(Y/X) \propto P(X|Y) \cdot P(Y)$$

Bayesian Classifier

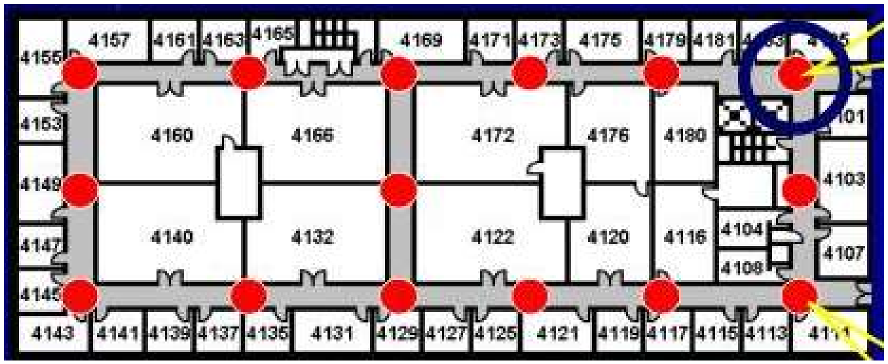
- Suppose, for a given instance of X (say x = $(X_1 = x_1)$ and ….. $(X_n = x_n)$).
- There are any two class conditional probabilities namely $P(Y = y_i | X = x)$ and $P(Y = y_j | X = x)$.
- If $P(Y = y_i | X = x) > P(Y = y_j | X = x)$, then we say that y_i is more stronger than y_j for the instance $X = x$.
- The strongest y_i is the classification for the instance $X = x$.

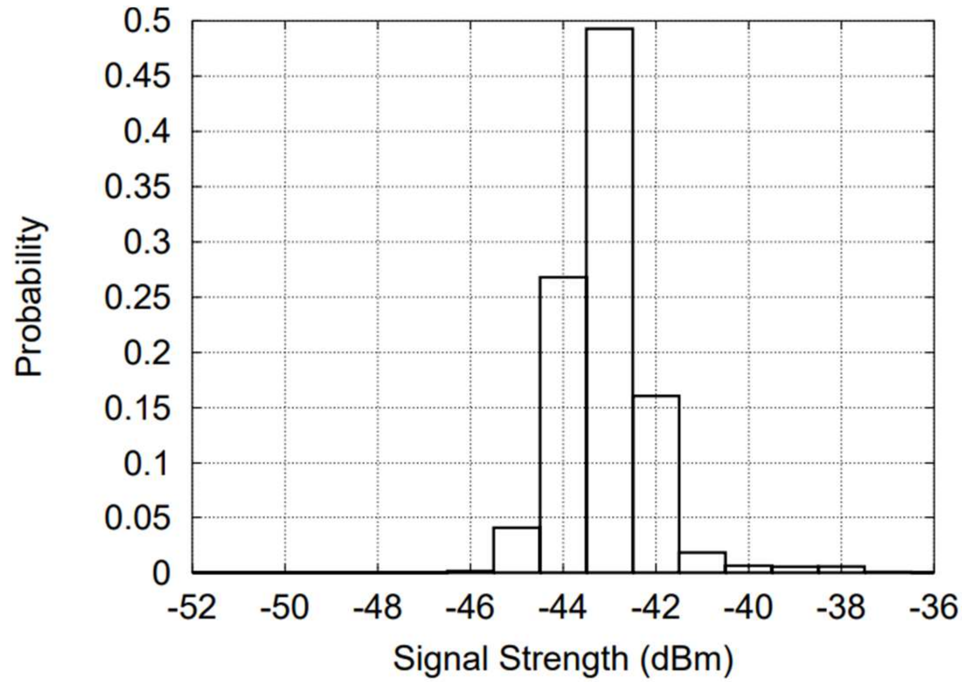
HORUS Localization: A Probabilistic Approach



https://www.slideserve.com/emele/horus-a-wlan-based-indoor-location-determination-system

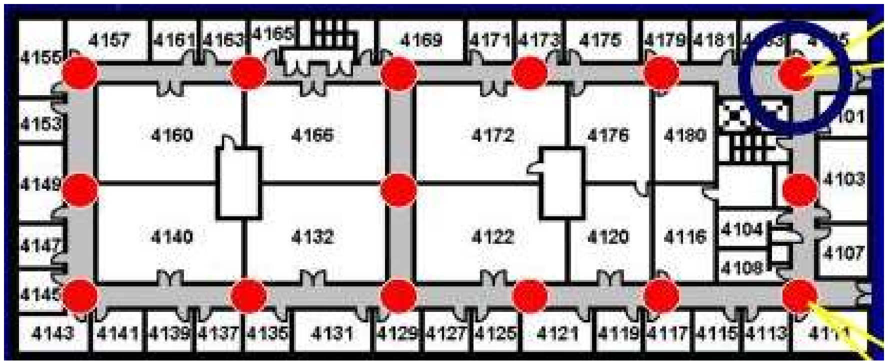
HORUS: Offline Phase

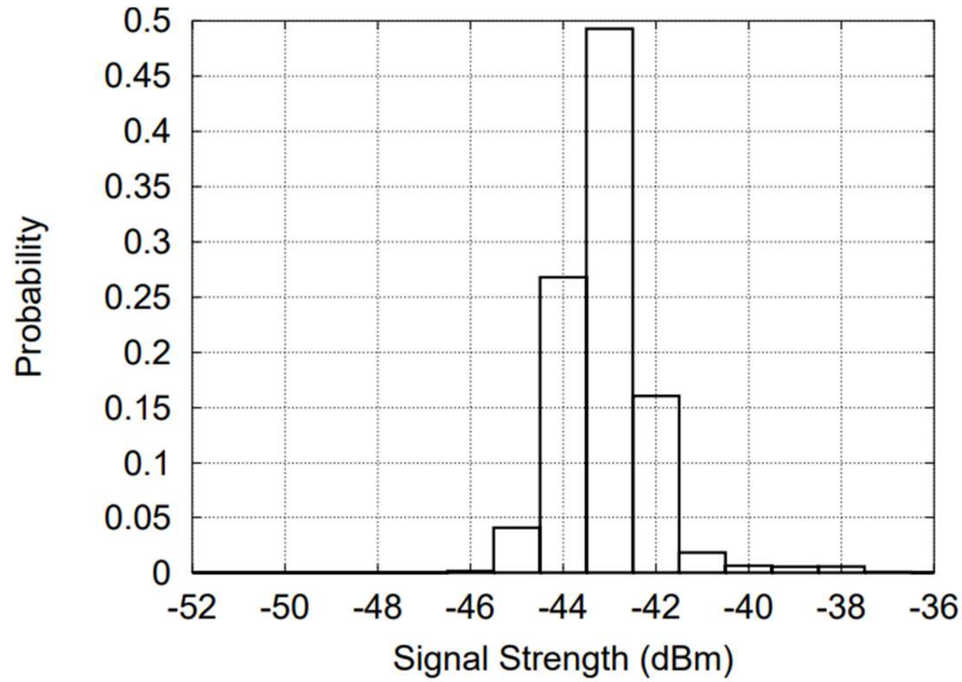




- Stores distribution of RSSI at different locations:
  - For location L, store: $P(RSSI \mid L)$

HORUS: Online Phase



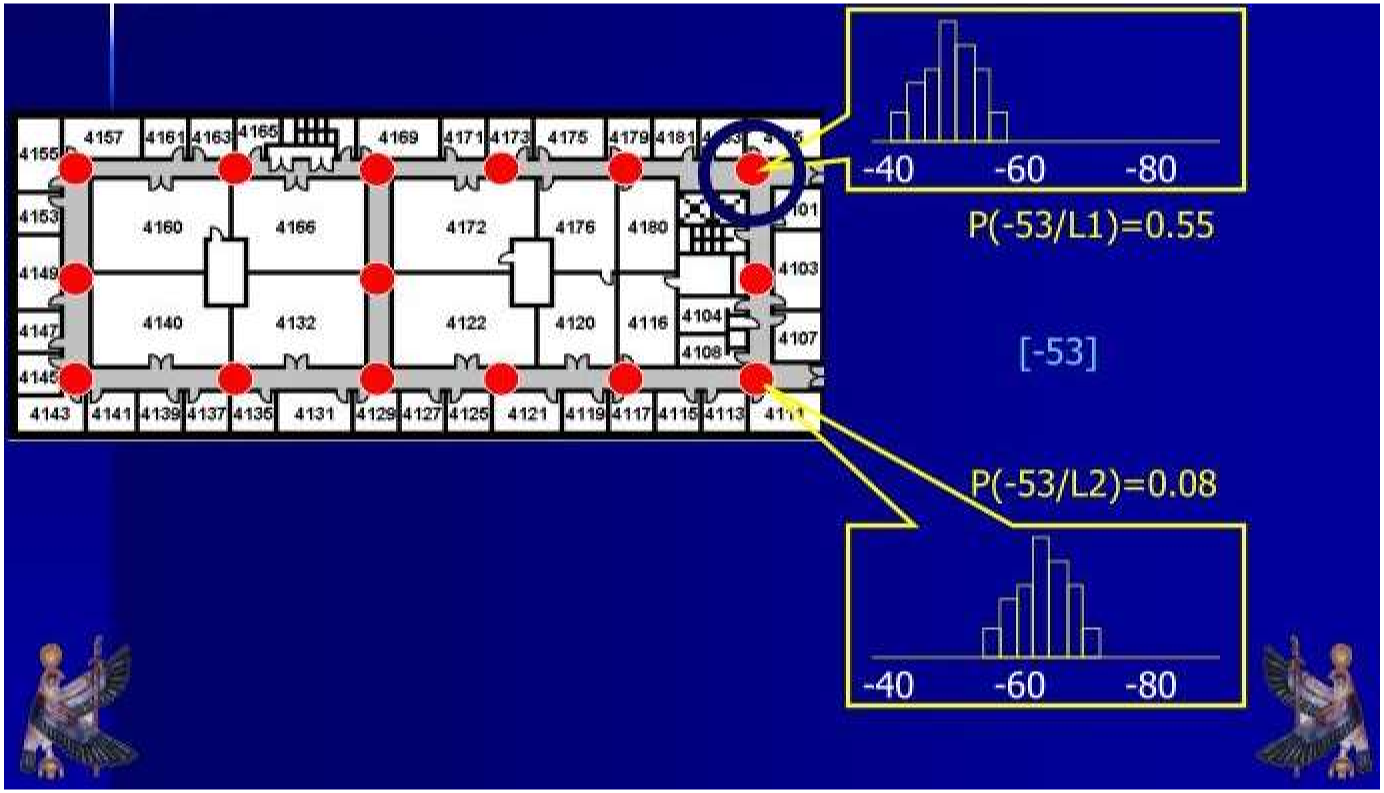


- Record the RSSI values from all APs (suppose there are k APs):
  - For all locations L_j, calculate:

$$P(L_j|RSSI) \propto P(RSSI|L_j) \cdot P(L_j) \propto P(RSSI|L_j) = \prod_{i=1}^{k} P(RSSI_i|\, L_j)$$

- The location L_j with maximum $P(L_j|RSSI)$ is more likely the device location

HORUS: An Example



Observed RSSI: -53 dBm

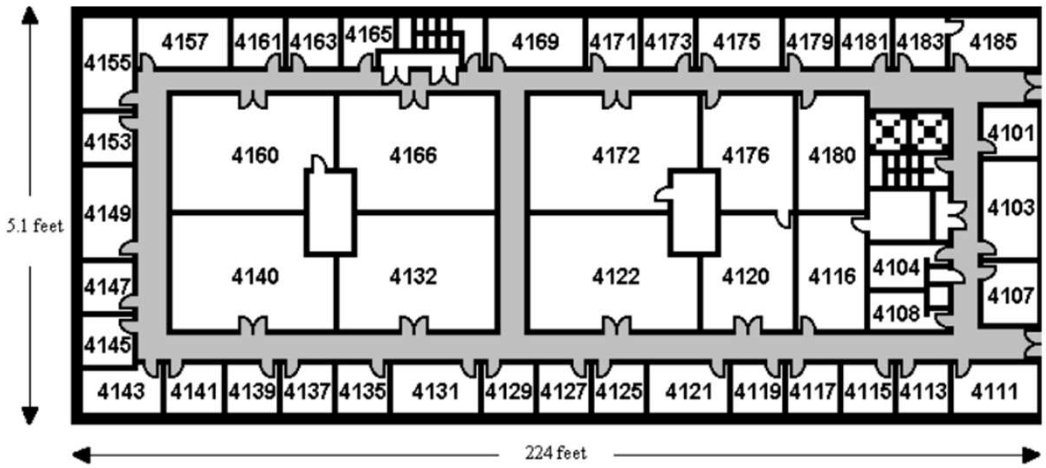
$P(L1|-53) \propto P(-53|L1) = 0.55$

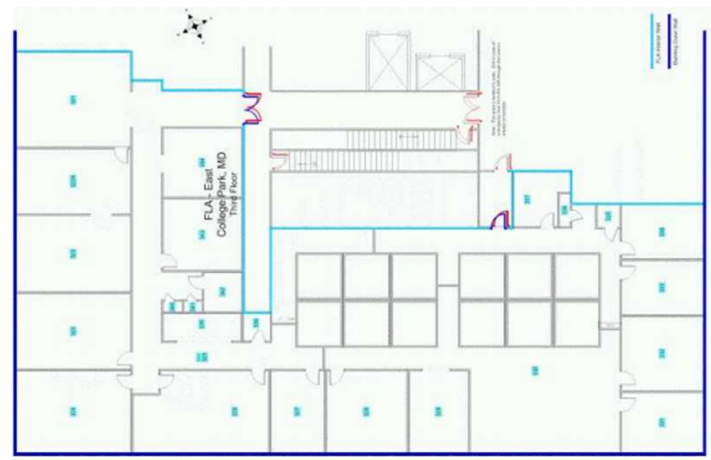
$P(L2|-53) \propto P(-53|L2) = 0.08$

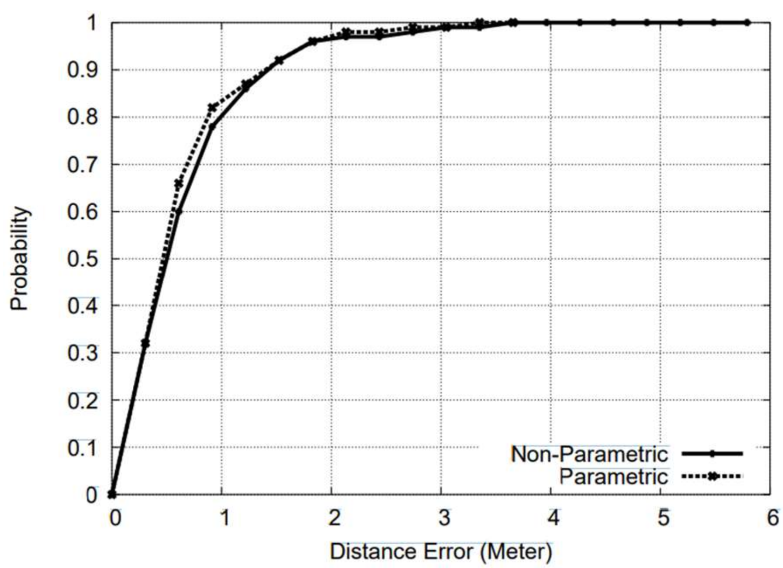
∴ The user is more likely in L1

https://www.slideserve.com/emele/horus-a-wlan-based-indoor-location-determination-system

HORUS: Experimental results





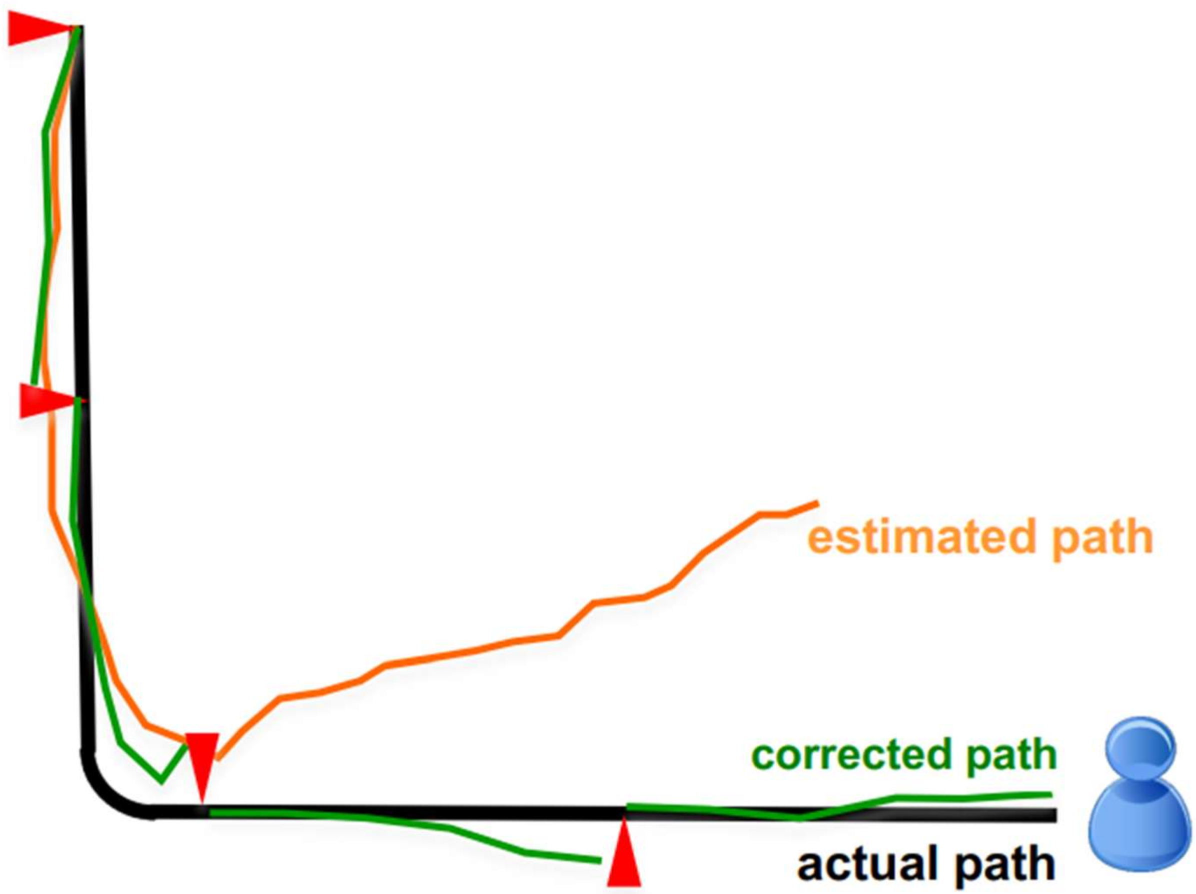


90th percentile error: 1.5 meters

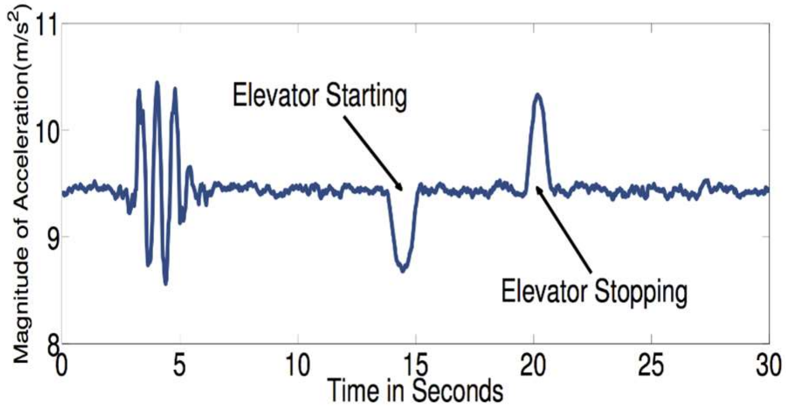
Indoor Localization: Unsupervised Learning Based

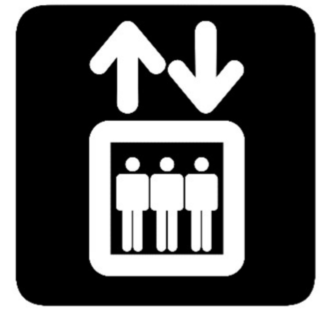
Src: R. Roy Choudhury, UIUC 2021

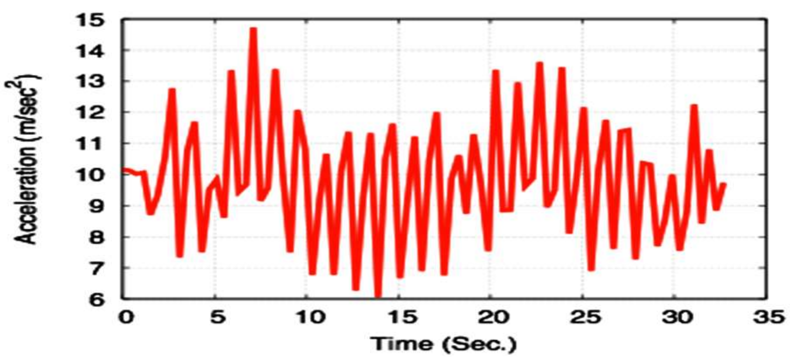
Path Estimation

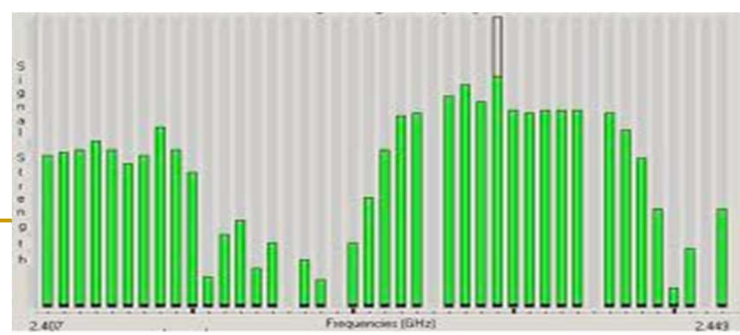


Many Many Landmarks



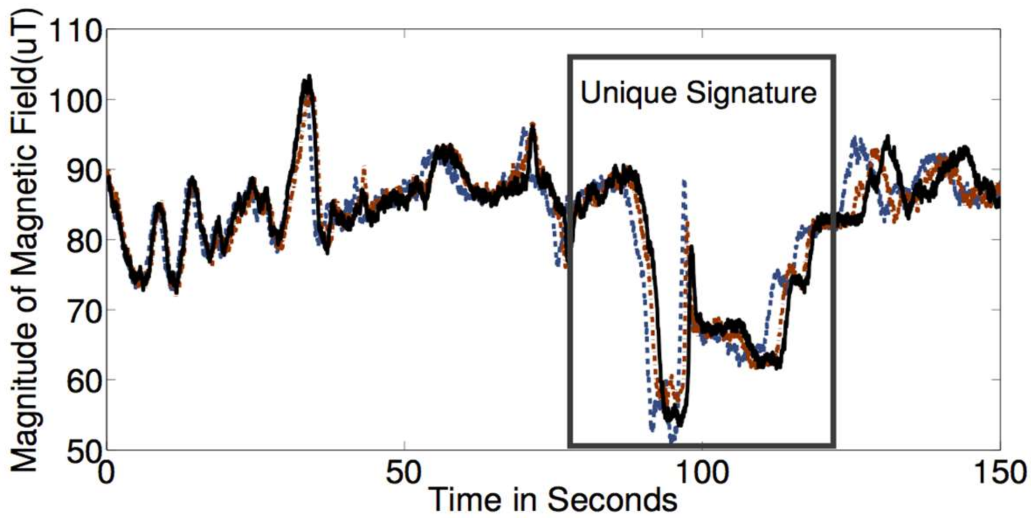






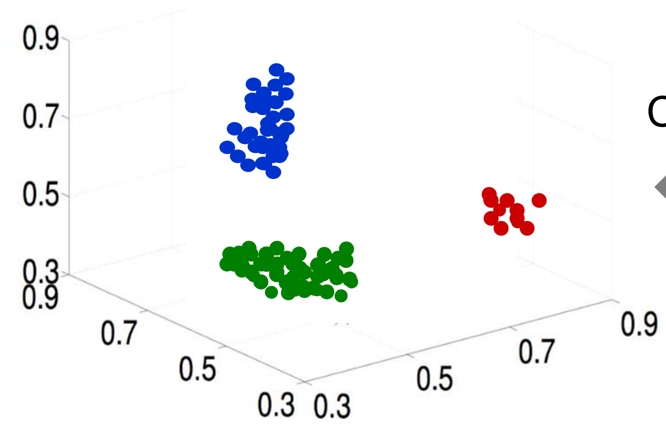
Let the Patterns Emerge

- Of course, not looking for pre-determined patterns
  - Rather, let patterns emerge from sensor readings
  - E.g., through multi-dimensional clustering



Raw Sensor → Landmarks

sensor space

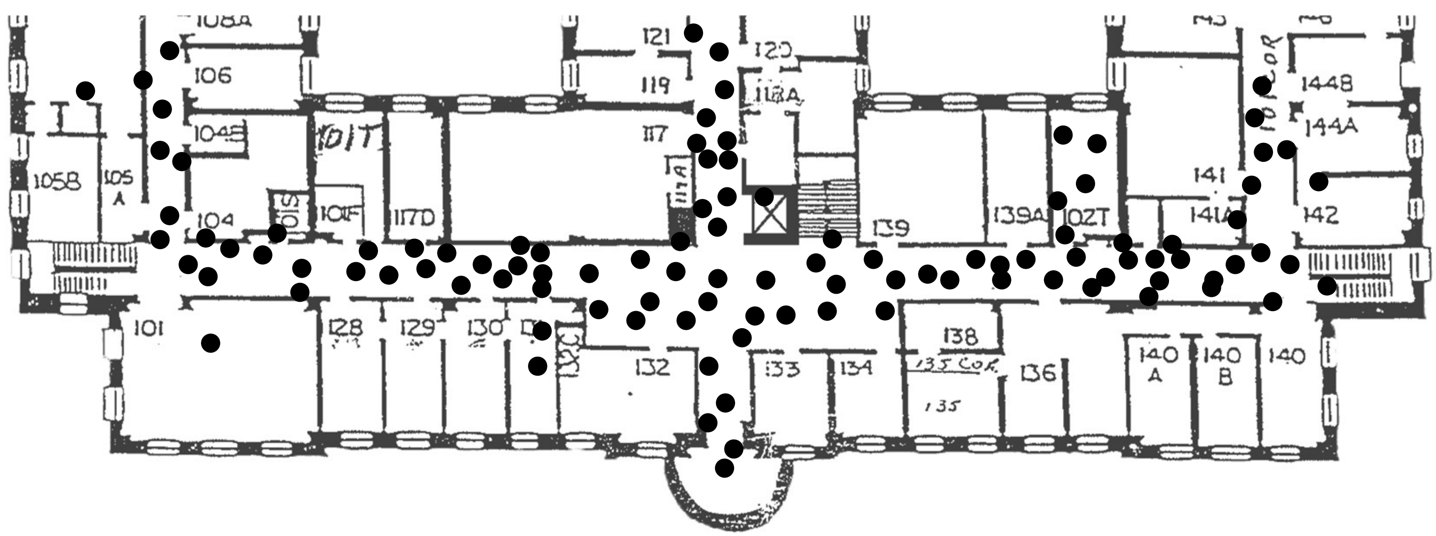


Clustering

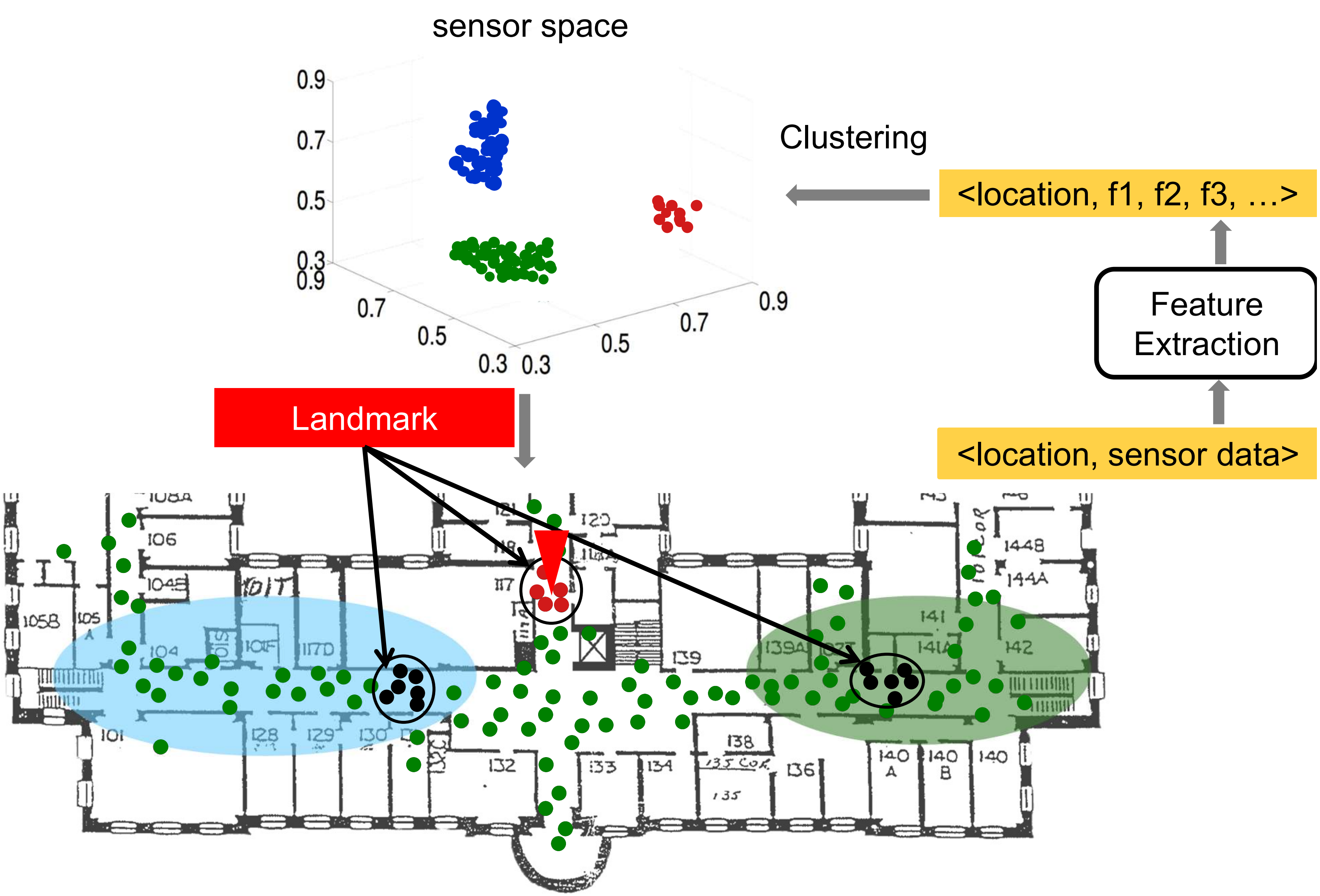
<location, f1, f2, f3, …>

Feature Extraction

<location, sensor data>

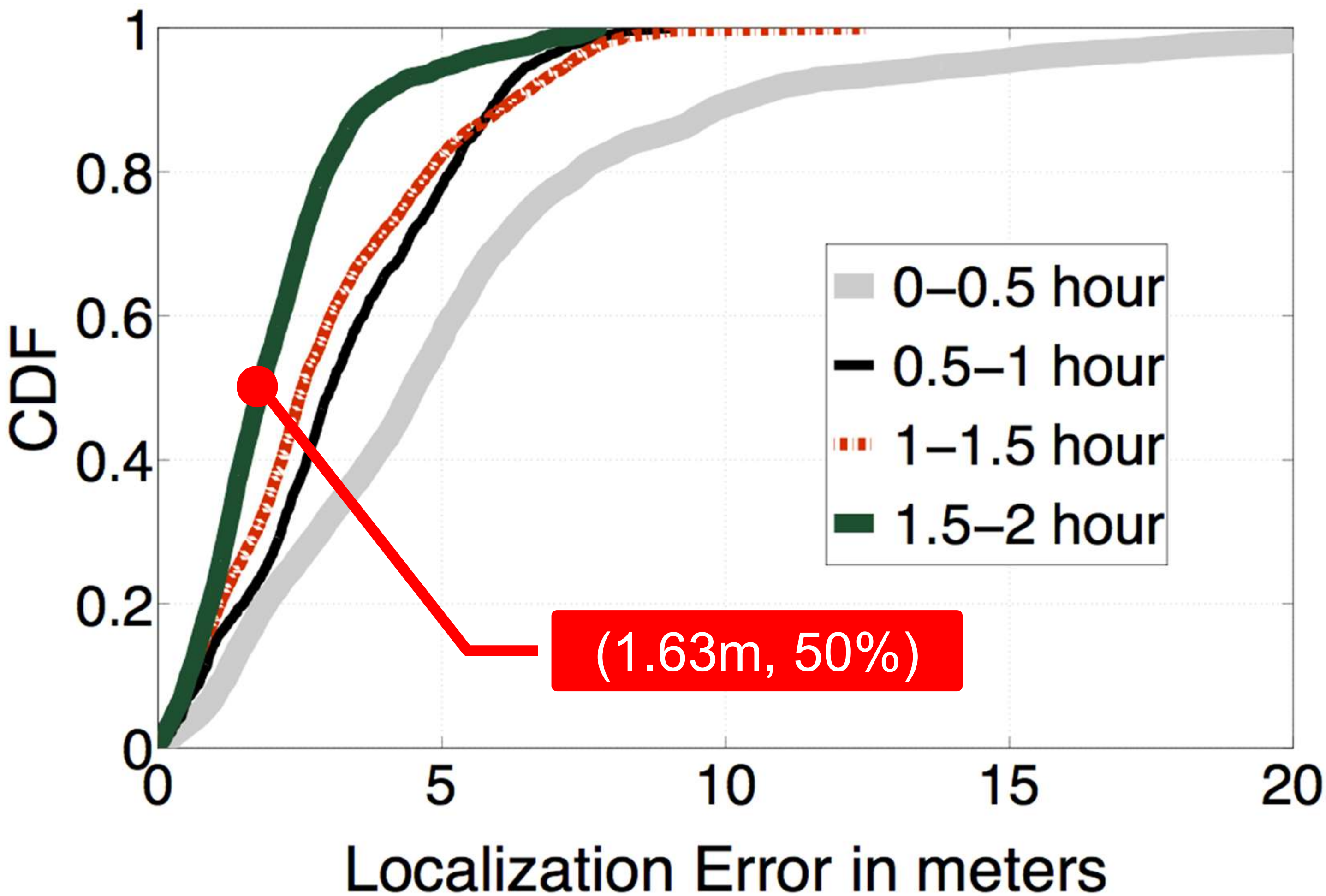


Raw Sensor → Landmarks



Performance

Experimentation on 8000 sq. meters



References

- P. Bahl and V. N. Padmanabhan, “RADAR: an in-building RF-based user location and tracking system," IEEE INFOCOM 2000, pp. 775-784.

- Moustafa Youssef, Ashok K. Agrawala, “The Horus WLAN location determination system”, MobiSys 2005, pp. 205-218.

- He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, Romit Roy Choudhury, “No need to war-drive: unsupervised indoor localization”, MobiSys 2012, pp. 197-210.